Attending to message and medium : the effects of planning time on the task-based language performance of native and non-native speakers.

Foster, Pauline Mary

Attending to message and medium:
The effects of planning time on the task-based language performance of native and non-native speakers

Pauline Mary Foster
King's College
University of London

Thesis submitted for the degree of Doctor of Philosophy
January 2000

Abstract

Speaking a language involves simultaneously attending to the message and the medium. A central concern of this thesis is to explore the way the brain allocates attention to these two demands, and how far they can be said to compete for limited resources. It has been argued that complex or unfamiliar subject matter requires relatively more of the speaker's attention. For native speakers this may have little impact on language form, as first languages demand relatively little attention to execute. But for second language learners difficult subject matter may mean they pay insufficient attention to language form, with the result that some target language forms might be slowly or poorly executed or even avoided altogether. Pedagogically, this would be an undesirable state of affairs.

This thesis examines the idea that attention to language form can be increased through allowing a person time to plan before speaking. Native and non-native speakers of English were recorded doing language tasks under planned or unplanned conditions. The results showed that both types of speaker were affected by the planning condition: pre-task planning resulted in language that was more fluent, more syntactically complex and (for the non-native speakers) more grammatically accurate. It is concluded that these performance gains were the result of increased availability of attention.

A further analysis revealed that the much greater fluency and complexity of the native speakers' language was due in part the use of prefabricated and memorised sequences of words, and that planning time reduced their incidence. However, the non-native speakers appeared to use many fewer memorised sequences and planning time had no measurable effect on this.

The implications of these results for second language pedagogy are discussed, and it is suggested that the goal of being native-like in using a second language is not always the same as being accurate or complex.

Acknowledgements

First of all, my thanks are due to Professor Peter Skehan, without whose unflagging encouragement and advice this study would never have been undertaken or completed. Thanks are also due to the Centre for Applied Linguistic Research at Thames Valley University for financial support during the final two years of my studentship. I am further indebted to Michael Foster, Kevin Germaine, Rosina Marquez, Jane Willis, Dave Willis, Peter Skehan, Alex Teasedale, Benny Teasedale and Alan Tonkyn for providing the native speaker intuition which enabled me to undertake an important lexical analysis of the data, and to the students at St. Mary's University College and Richmond Adult and Community College who allowed me to record their voices.

Finally, I acknowledge the patient support of my family, especially my children, who for many years have watched me struggle with the letters S A, M A, and PHD. I am glad to tell them that there are no more letters to do.

Table of Contents

List of tables and figures

Chapter One

Introduction and overview

The thesis presented here is concerned with language learning and language use, a broad field of study, ploughed by many different lines of research. A good starting point therefore for this introduction is to set out which lines I do not attempt to follow, not because I believe them to be unproductive or unimportant, but because they are not immediately connected with the cognitive perspective that I will take.

For example, there is no consideration given in this thesis to the way second language acquisition may be influenced by age, gender, or individual differences in motivation, aptitude or learning style. Nor is there any concern with pragmatic, ethnographic or sociocultural issues. The impact of a person's first language on the process of learning a second is discussed in the literature review in chapter three, but does not figure in the research study itself. Similarly, there is no concern in the research study with linguistic universals and the biological basis for first language learning, though this too is the subject of some discussion in chapter two. Finally, the importance of interaction in second language acquisition, which has been the focus of a great deal of research interest in recent years, is discussed in chapter four but not pursued as a productive line of inquiry.

This research study explores the idea that, although first language learning may be a special case, second language learning is not qualitatively different from other kinds of learning. Thus the emphasis here is not on the nature of language itself as a special, biologically programmed endowment, freely available to children and reactivated with difficulty by adults. Instead the emphasis is more generally on the nature of learning a complex skill, and what can be discovered about the cognitive processes that underpin its

fluent execution. From this perspective, second language learners are seen as operating a complex processing system that deals with linguistic information in the same way it deals with other kinds of information. The interest is less in the linguistic knowledge which learners may have acquired than in how they develop fluency and accuracy in accessing it, and what strategies they might employ when their knowledge is inadequate or inaccessible to the demands of real time performance. Ultimately, the aim for such an approach is to illuminate the cognitive processes involved in the learning of complex skills to see what pedagogical implications might be drawn for the language classroom.

This research study is thus firmly located in cognitive psychology and what is known about the way human beings acquire complex skills. Accordingly, chapter two looks in some depth at the work of cognitive psychologists over the last twenty years or so into what has become known as information processing. The early work in this field (Shiffrin and Schneider 1977) was not done with the particular skills of producing and understanding language in mind, but rather in the way people could be trained to respond to certain sets or sequences of digits and letters. Nevertheless, this research is dwelt on at some length because it gave rise to the notions of *controlled* and *automatic processing*, and a conceptualisation of learning that involves a shift from the former (a slow and effortful application of knowledge) to the latter (a faster, effortless application of knowledge) through repeated practice. Anderson (1982), using the differently named but conceptually similar labels of *declarative* and *procedural* knowledge, sought to bring first language and second language learning into this model, and the second section of chapter two examines how justified this application is.

An important feature of the model of the brain as an information processor is that its capacity for dealing with controlled/declarative knowledge is constrained by attentional space. Because of this processing limitation, the brain cannot carry out two simultaneous tasks requiring controlled processing (such as drawing a circle with one hand and a square with the other) without disastrous loss of performance (i.e. the halting execution of two very irregular shapes). In order to be successful at performing two or more tasks at once,

the brain must be able to use automatic processing of procedural knowledge to ease the processing load. In this regard language can be seen as something which normally requires the brain to attend to two things at once: the content of the message and the form in which it is encoded. The middle sections of chapter two look at different accounts of how the brain copes with this burden in first language performance. For Ochs (1979) and Givon (1979) the strategy is one of avoidance: the production of complex syntax requires more processing capacity than the production of simple syntax so when a situation or the subject-matter is particularly attention-demanding a speaker will not attempt complex syntax but rely instead on simpler, less attention-consuming constructions. An alternative strategy is proposed by researchers such as Pawley and Syder (1983) who note that native speaker do not exploit the grammar and lexis of their language to anything like their full potential, preferring instead to use words in the same conventional sequences as everyone else. Native speakers have a store of many thousands of such sequences and draw on them regularly when encoding and decoding language, thus saving themselves the burden of constantly processing language from scratch. In this view a native speaker produces utterances not so much through proceduralised knowledge of syntactic rules, but by cobbling together ready-made phrases. This would account, for example, for the way people having to produce language in real time and under some communicative stress (as in a live broadcast) are apt to combine unanalysed chunks with unintended and hilarious results.

> Some of these players never dreamed they'd be playing in a Cup Final at Wembley, but here today their dreams have come true.
>
> We all have a chip on our shoulder that we want to get off of our chest
>
> Yesterday he played the final card in what has been a tricky game of chess

Fantoni (1985)

These strategic short-cuts are further discussed in chapter two in relation to Ellis' (1996) account of first language acquisition as a process of committing to memory regularly encountered sequences from which lexical and syntactic rules are ultimately extracted. The use of pre-formed 'lexicalised' sequences can thus be seen as both a learning and processing strategy. The extent to which they are used by second language learners is a subject pursued further in the research study itself and examined in considerable detail in chapter eight.

It is remarked in chapter two that the notions of information processing are not applied as commonly to native speaker language as they are to non-native speaker language. Chapter three, therefore, reviews the relevant research which has sprung up in the field of second language acquisition. Central to the debate is the nature of attention itself, and the extent to which learning is dependent on the brain's conscious or unconscious attention. Related to this are the roles of implicit and explicit knowledge in second language learning, in particular whether there is interference or interface between them. Research from cognitive psychology is discussed in which it is shown that structural learning can be shown to operate independently of conscious awareness or conscious will. For Krashen (1981, 1982, 1985) this is the only way for successful language learning to proceed. Other research, however, into the learning of both artificial and natural languages shows that explicit knowledge can in fact facilitate the development of faster and more accurate (automatic/proceduralised) skills, pointing up a role for explicit teaching of language rules.

The practice of second language teaching is reviewed at some length in the following two chapters, from the behavioural conditioning of pattern practice to the more modern attempts to contrive an implicit focus-on-form in the classroom. At each turn the pedagogical assumptions on the nature of language learning are examined in the light of what we know about information processing and the development of automaticised/proceduralised knowledge. Of central importance throughout is the notion of the brain as having limited capacity to process information, so that too much attention

to one aspect of the target language, e.g. the message it is conveying, will take away attention from another aspect, e.g. the form in which the message is couched. Because the message is usually more important and interesting than the form (indeed the attending closely to the form of the message is often unnecessary in understanding the meaning) it is a challenge for pedagogy to bring about a focus on form without reducing the second language learning to the mastery of an unproductive system of rules. (Doughty and Williams 1998)

Task-based approaches to second language teaching and learning are given special consideration here because the research study is located within Skehan's (1996) framework for task-based learning. In this, the design of a task, its content and its implementation are all possible ways for the teacher to manipulate the learner's attention between the primary focus on the meaning of what they are doing and a secondary focus on language form. The limitations of a learner's attention are an important consideration. If the task is cognitively taxing because, for example, its subject matter is perplexing or complex, it is more likely that the learner will have insufficient attention during the execution of the task to give to formal aspects of the language. Accordingly, in chapter five research is surveyed into factors that contribute to the cognitive load of a task and how these have been shown to impact on various aspects of second language performance. Another line of inquiry into helping a learner's attention encompass both form and meaning during tasks has centered around giving planning time before the task begins, on the assumption that this will ease the processing burden and increase attentional capacity. Research evidence from these investigations has pointed to significant gains in fluency and complexity in the language produced by learners who were allowed planning time before attempting the task.

The research study described in chapters six and seven of this thesis is an investigation of the interaction of these two influences of planning time and task complexity on the language produced during tasks. Three language features are examined: how fluently delivered, grammatically accurate and syntactically complex language is under planned and

unplanned conditions. Uniquely, the study encompasses both native and non-native subjects. This is to connect more robustly the research in cognitive psychology, which typically uses native speaker subjects, with research into language learning, which typically uses non-native speaker subjects. One of the most important research aims is to examine the extent to which speakers using their native language are affected by the cognitive demands of the tasks they are set, what coping strategies they use, and whether these are changed by giving planning time. It is against this 'baseline' data that we can more properly assess what the non-native speakers, with their added burden of having to process an imperfectly learned language, are able to do with tasks of varying cognitive difficulty under different planning conditions. In chapters eight and nine the results of the study are presented in detail and discussed in the light of related research studies, some of which preceeded this one, though others appear to have been been carried out simultaneously. It is gratifying to discover that investigating the nature of planned language and task complexity is such an active and productive area.

Ultimately, the aims of the research study were pedagogic and I would hope that the conclusions I draw here in chapter nine can illuminate choices made by second language teachers and oral examiners. A corollary to this investigation has been an unexpected illumination of the shortcomings of some of the tools used in empirical research into second language acquisition. During the course of this study I worked with fellow researchers to improve one of these, a unit to measure spoken interaction, which I explain briefly in appendix D. This was a long, hard but necessary detour and it is my hope that in addition to benefitting second language teachers, the outcome of this study will benefit researchers toiling in the same field.

Chapter Two

The psychology of controlled and automatic processing.

Over the last twenty years there has been a considerable amount of work in experimental psychology into the way the human brain processes information. From this has emerged a considerable degree of agreement as to the likely nature of this processing. Two main models have been proposed: the ACT* models of Anderson (1982, 1983, 1992) and the controlled/automatic processing model of Shiffrin and Schneider (1977). Linguists interested in the way the brain processes language refer with remarkable frequency to one or both of these models in reports of their own research. It is important therefore to look at them in some detail before moving on to assess how far they can account for many of the observed characteristics of first and second language performance. The chapter explores how language processing strategies are thought to influence the way speakers perform under time pressure, and how, ultimately, this can account for the evolution of syntax itself.

2.1 Controlled and automatice processing

In 1977 Richard Schiffrin and Walter Schneider undertook a series of experiments that have since become classics, cited frequently and with approval in both psychology and linguistic research. In the first experiment of the series, (Schneider and Shiffrin, 1977) subjects were given one to four items to remember. These were either all digits or all consonant letters of the alphabet and were termed the *memory set*. The subjects were then shown in rapid succession twenty cards (each called a *frame*) with one to four digits or letters on them, and were asked to identify any that contained an item from the memory set. A memory set item that appeared in a frame was called a *target,* a non-memory set item that appeared in a frame was called a *distractor*. Subjects were given over two thousand *trials* of twenty frames each to search for targets, with a fresh memory set to learn before every trial. The difficulty of the task was manipulated by increasing the *load*,

(the number of items in the memory set multiplied by the number of items in each frame), and also by manipulating the number of milliseconds each frame was shown.

The most important factor in this experiment was the way in which the memory set items were related to the distractors. In the condition known as consistent mapping (CM) across all trials the memory set items were never used as distractors, and distractors were never used as memory set items. Memory set items were always digits, and the distractors were always consonant letters of the alphabet. In the condition known as varied mapping (VM) however, memory set items and distractors were all from the same category, (consonant letters of the alphabet), and were randomly intermixed over the trials. A consonant letter which had appeared in the memory set for one trial could therefore act as a distractor in another trial.

Success at the tasks was measured by the percentage of correctly identified targets *(hits)* at each frame time from 800 msec to 40 msec. The results showed that in the CM condition the percentage of hits was not affected by load (the number of items to remember multiplied by the number of items to search) and only a little affected by the length of time frame. In the VM condition, however, the percentage of hits was strongly affected by load and frame time and was worse than any of the CM results. To Schneider and Shiffren these results suggested that in the VM condition subjects had to keep checking the items on each frame against those in the memory set. The heavier the load, the less successful the search, indicating that this process was *capacity-demanding* (i.e. needed lots of attention). The CM condition however enabled subjects to develop an automatic detection process which was not compromised by load and was thus *capacity-free* (i.e. required no attention),

A subsequent experiment explored the speed with which subjects correctly identified target items in a frame. The experiment used a very similar design to the first except that only one frame was shown per trial and subjects' reaction time was measured. It was found that under VM reaction times were very significantly slower. These results

supported the conclusion drawn from the first experiment: the process by which the brain detects a target under CM conditions is different from the process it uses to detect a target under VM conditions. Shiffrin and Schneider called these modes *automatic information processing* and *controlled information processing.* They argued that the brain, when confronted by an unfamiliar search-and-detect task, uses a comparison process whereby each item in each frame is checked against the memory-set items to see which are targets and which are distractors. In the CM condition, where the same items are either always targets or always distractors, the brain can bypass this item-by-item search by gradually developing an automatic response that is faster and more accurate. In the VM condition however, because the same items are both targets and distractors, the brain can not develop an automatic response and continues to use a controlled search.

Further experiments were designed to test the nature of automatic information processing and how it develops from controlled information processing. The first of these (Shiffrin and Schneider 1977) used only CM in a design otherwise identical to the search-and-detect studies outlined above, with the important difference that the memory-set and distractors were now composed exclusively of letters of the alphabet. Shiffrin and Schneider had realised that by having memory-set items and distractors in distinct categories (i.e. letters and digits) they had possibly provided an unfair advantage to the CM condition that could have accounted (at least in part) for the rapid development of automatic responses. It was felt that this advantage should be removed, so that any development of automatic responses would now be without the benefit of previously-learned categories. Subjects had to learn a memory set of two letters (selected randomly from an ensemble of nine) and to detect targets in a series of 1,500 trials at a frame time of 200 msec. For the following 600 trials this frame time was reduced to 120 msec. At this point the memory ensemble was switched to a different nine letters that had hitherto been used as distractors. A further 2,400 trials were run with new memory sets at a frame time of 200 msec. The percentage of targets correctly identified was recorded, as was the reaction time.

The results were extremely interesting. The hit rate for the first 1,500 trials rose from 50% to 90%, whereas falsely identified targets fell from 12% to 3%. Reaction time for identifying hits quickened from 770 msecs to 670 msec. The shift to a faster frame time for the next 600 trials caused some loss of performance but this was soon restored to a peak hit rate of about 82% and false alarms steadied at 5%. The authors conclude that the subjects were using controlled searches at the beginning of the experiment, but by the end of 2,100 trials had developed an automatic response to the memory ensemble of letters This conclusion was reinforced by the subjects themselves who reported that they were aware at the beginning of the experiment of the need consciously to rehearse the memory set as they scanned each frame for targets. After about 600 trials however they were no longer aware of any such demand on their attention. The fact that the memory and distractor sets were both composed of letters did not prevent the acquisition of an automatic response, suggesting that both categories and automatic responses can and will be learned for arbitrary sets of easily-confused memory items.

Shiffrin and Schneider had hypothesised that if an automatic response had been learned for a particular stimulus, it should be very hard for the subject to change it or *un*learn it in the short term. This is what the second part of the experiment was designed to test. Having learned an automatic response to a particular set of nine letters over 2,100 trials, subjects had to switch to a different memory ensemble of the nine letters they had come to recognise as distractors. The results were a strong endorsement of the hypothesis. Immediately after the switch, the hit rate plummeted to below the 50% (i.e. chance) level it had been at the beginning of the experiment when subjects were completely unpracticed in target detection. It took 2,400 trials with the new memory ensemble before they were performing again at the 90% hit rate achieved after the initial 1,500 trials. Shiffrin and Schneider conclude that the automatic response developed for the first memory ensemble resulted in a negative transfer that hindered the development of a new automatic response to the second memory ensemble, and this is why the second automatic response took so much longer to become established.

This experiment demonstrates the long-term nature of automatic processing. An automatic response can take thousands of trials to become established, especially when the brain has no previously learned separate categories (letters, digits) in which to put the targets and distractors. But once these automatic responses have been acquired they are difficult to suppress, and require thousands more trials to dislodge.

Further experiments reinforced the conclusion that when an automatic response has been established it is not under the conscious control of the subject. To test this, subjects had first to develop an automatic response to memory-set items under CM conditions, and then were instructed to identify targets only on the top left to bottom right diagonal in a frame, ignoring targets that appeared on the other diagonal. The results showed a falling off in performance from a hit rate of 84% when all targets were to be identified, to a rate of 62% when only targets on a particular diagonal were to be identified. In order to identify targets appearing on the valid diagonal subjects had to substitute a controlled search for the automatic response they had built up through CM training. The subjects showed however that they could not always ignore targets that appeared on the invalid diagonal *even when they knew these would not be hits.*[1] When a similar study was carried out under VM conditions to see if subjects could ignore targets on an invalid diagonal, it was found that such targets did not cause any impairment to the hit rate. Because subjects were using a controlled search process (resulting from the VM training) they were able to focus their attention only on the valid diagonal.

Summing up what Shiffrin and Schneider conclude from their studies, we can say that information processing has two fundamental modes that differ considerably. Automatic processing needs a lot of training to acquire, cannot be modified without considerable effort, takes place without the subject's conscious attention and even against the subject's

---

[1] The strength of an automatic response and the difficulty of suppressing it was further illustrated by subjects' reporting that when they attempted to read anything after an experimental session, the letters they had learned as consistently mapped memory-set items would 'jump out' of the page at them. (Shiffrin and Schneider 1977.153)

will, is rapid, efficient and capacity-free. Controlled information processing, on the other hand, requires no training, can be adopted and adapted easily, is under the conscious control of the subject, is slower, less efficient and capacity-demanding.

From these results Shiffrin and Schneider propose a general theory of learning. They conceive of memory as a very large collection of nodes, each representing a grouping or set of informational elements, that become 'complexly and increasingly interassociated and interrelated through learning' (p154). Most nodes are normally inactive, and when in this state are called the *long-term store* (LTS). This is our permanent and passive repository of learned sequences and items of information. In the presence of some external stimulus a group of nodes become active and in this state are called the *short-term store* (STS). This is a temporary repository from which any freshly received information may immediately be lost or forgotten when the external stimulus ends. Automatic and controlled processing are the two ways information moves into and out of the STS. An automatic process is a learned sequence of nodes from the LTS that is always activated and moved into the STS in response to a particular stimulus, without any conscious attention or control by the subject. An example would be reaching up to a light switch when entering a dark room. This movement does not have to be consciously chosen or guided. If you have entered the same room many hundreds of times your hand will rise to exactly the right height and in exactly the right position for your finger to make contact with the switch without your conscious control or guidance of any of the muscles involved. Even when you know the bulb has blown and the light will not work, it is difficult to prevent your hand rising to the switch in automatic response to the stimulus of entering the dark room. And anyone who has ever driven a friend's car where the controls are in a different position to their own will have experienced the frustration of repeatedly activating the windscreen wipers when meaning to signal a turn.

Controlled processes activate a temporary sequence of nodes in the LTS and move these into the STS. They always require the attention and control of the subject in order to be executed. Because these temporary sequences require attention, only one at a time can be

activated without causing interference between them. Controlled processes are thus severely capacity-demanding, but they have the important benefit of being easy to alter. Novices at the piano find it very hard to play a new piece with both hands at the same time, needing to attend to each hand separately, but perceived errors in fingering, notes or rhythm can be suppressed and replaced with relative ease.

In this model, learning involves the replacement of controlled processes with automatic processes as information is eventually transfered from the temporary repository of the STS to the more permanent repository of the LTS. The mechanism responsible for this transfer is the *strengthening* of the connection between the nodes as a result of repeated exposure and practice. This model does not assume that the structure of the response is modified, only that it is executed more rapidly. Learning takes time, but once an automatic response has been developed it remains as a fixed series of interrelated nodes in the LTS that can be triggered without any loss of capacity in the STS. Shiffren and Schneider see this process as fundamental to the acquisition of a complex skill such as reading. The lower order skill of letter-recognition is first a controlled process requiring most if not all of the capacity of the STS. Through practice it becomes automatised and established in the LTS freeing up attentional 'space' in the STS that can be devoted to the higher order skill of word recognition. This in turn will become automatised and allow the STS to allocate attention to the meaning of phrases or whole sentences. As previously taxing components become progressively automatised and capacity-free, a slow and difficult task becomes swift and easy. In this way, controlled processing lays down the 'stepping stones' that allow the learner to progress to more complex levels of processing.

The work of Shiffrin and Schneider reinforces the findings of other researchers in the area (e.g. LaBerge and Samuels, (1974) on the skill of reading) and is frequently cited with approval. In one important respect, however, it has been criticised by Cheng (1985). She says that the dichotomy between capacity-demanding and capacity-free processing is neat and intuitively appealing, but wrong. She challenges the conclusion reached by Shiffrin and Schneider that automatic processing is always a capacity-free speeding up of a capacity-

demanding process, citing a study by Hoffman, Nelson and Houck (1983) which showed that when subjects were asked to perform two CM search-and-detect tasks together, there was evidence of significant interference from one task to another, suggesting that the detection of CM targets can not therefore be capacity-free. Cheng maintains that many of the results obtained by Shiffrin, Schneider and colleagues (e.g. Fisk and Schneider 1983), which they report as evidence for the development of automaticity, could be accounted for differently. She points to many instances in their data where an increase in response time and hit rate could be explained by subjects developing a *category strategy* to distinguish targets from distractors. This is most obvious, as Shiffrin and Schneider themselves concede, when targets are all digits and distractors are all letters. Subjects do not have to invent categories for these: they already exist. But even when targets and distractors are made up from the same pre-existing category, it is quite possible for subjects to develop new categorical distinctions between the two sets that will increase the speed at which they can be identified. This is not the same as automatising a controlled response: it is *restructuring* the response so that it becomes more efficient. Cheng says a similar analysis would apply to a task such as finding the sum of ten 2s. A person who knows how to add up but not how to multiply would do the same addition function nine times over in order to get the answer. A person who knew how to multiply would do just one multiplication function to get the answer. It is obvious, however, that the multiplication function is not an automatised serial addition, but a restructured solution to the task that brings an impressive gain in speed and accuracy.

Cheng cites other research to support her analysis. In one such study, Hirst, Spelke, Reaves, Caharack and Neisser (1980), subjects had to perform two complex tasks simultaneously, i.e. playing the piano and repeating a heard sentence. After extensive practice, there was a marked improvement in performance which Hirst et al. interpret not as the development of automaticity (as each task was already automatised) but as the development of new patterns of response. In other words, the tasks had been restructured. Cheng's final example is offered as an unanswerable challenge to the simple dichotomy between automatic, capacity-free processing and controlled capacity-demanding

processing. It is easy for anyone to play three even notes on the piano, and easy to play four even notes. Performed in isolation, neither task is at all demanding of attention. To combine three even notes on one hand with four even notes on the other is, however, extremely taxing, demanding a lot of attention and repeated practice. Cheng concludes that it cannot involve the automatic and capacity-free execution of simultaneous tasks, but rather it is a new task requiring the development of a restructured skill.

Cheng's analysis does not completely invalidate the theory of learning postulated by Shiffrin and Schneider, but suggests that it is a much more complex process than the gradual transfer of sequences of nodes from the short-term memory store to the long-term memory store. There could be co-existing in the brain different kinds of learning which are useful for different kinds of skill. Learning a simple skill, like recognising target letters among distractors, or the position of the light switch in a room, may well be a straightforward move from controlled to automatic processing brought about by repeated practice, but the learning of complex skills cannot be accounted for in this way. Complex skills, such as playing the piano or speaking a second language, are learned through a process of restructuring that allows the learner to incorporate new information into that which has already been learned. This is a view of learning supported by other psychologists. For Rumelhart and Norman (1978) restructuring, or the imposition of a new organisation on information already stored, can account for the sudden moments of insight or leaps in proficiency that many language learners and piano players experience. Karmiloff-Smith (1986) agrees that restructuring is an important part of the learning process and results from the learner's 'drive' to incorporate hitherto isolated procedures into a 'unified representational framework'. Kolers and Roediger (1984) also see learning as an increase in control resulting from procedures in the memory of the learner undergoing what they call 'reassembly' and 'refinement'. In short, there is not necessarily a straightforward dichotomy between automatic (fixed, capacity-free, unconscious) processing and controlled (loose, capacity-demanding, conscious) processing, but more likely a continuum of automaticity, in which processes are open to readjustments which make them less capacity-demanding and more efficient.

Schneider and Detweiler (1988) have addressed many of these challenges to the Shiffrin and Schneider model and have proposed a revised model in which automatization is seen as a gradual transition from fully controlled to fully automatised processing through five identifiable phases, accepting that the processing of some skills may never reach totally capacity-free automatization. This is helpful in regard to language production, as it accounts, for example, for why otherwise completely fluent native speakers need to attend consciously to subject-verb agreement in complex utterances. But Schneider and Detweiler's model was not conceived specifically with language skills in mind, and indeed was not tested in the laboratory by any experiments that explored language skills so that its application to language processing, while potentially illuminating, remains speculative.

2.2 Declarative and procedural knowledge

Anderson's ACT* (ACT-Star: Adaptive Control of Thought) model of information processing (1982, 1983, 1992) is a broadly similar model of learning, although the terminology is different. And unlike Shiffrin and Schneider, who were more concerned with the characteristics of controlled and automatic processing than with how they develop, Anderson is interested in the details of the learning mechanisms that are responsible for automatization.

For Anderson, learning a skill is a cognitive process that can take many hundreds of hours. In the initial stages the learner, upon receiving instruction and knowledge about a particular skill, stores a set of facts about it which he can use in performing the skill. This is the *declarative stage*, because the facts have to be rehearsed in the learner's working memory (and commonly spoken out loud) for the skill to be performed. This slow and often very conscious application of declarative knowledge (roughly the equivalent of Schiffrin and Schneider's controlled processing) is gradually replaced through repeated practice by the *procedural stage* (roughly the equivalent of Schiffrin and Schneider's

automatic processing) in which a skill can be performed quickly and effectively without the help of verbal mediation or rehearsal in the working memory.

Anderson conceived his ACT model to explain how humans learn simple mathematics as well as complex algebra and geometry. These skills lend themselves very readily to being represented as a string of *productions* each of which is a necessary step or *goal* to the solution of a problem. An addition such as:

614
438
<u>683</u>

can be conceived as a series of twelve productions characterised as IF - THEN propositions that take you step-by-step through the adding up of each column and carrying digits over to the next. For example, the fourth production for this addition is described thus:

IF the goal is to iterate through the columns of
an addition problem,
and the last column has been processed,
and there is a carry,
THEN write out the carry,
and POP the goal.
(Anderson 1982:371)

The production system for solving an algebra problem is far more complex, but is basically the same: a sequence of goals which must be followed in order to arrive at the solution. Through the *compilation process* these productions gain in speed and accuracy as multiple steps are composited into single steps and the need to hold long-term memory information in working memory is eliminated. Adding up the digits in columns of figures and carrying over figures from one column to the next on the right becomes a fluently executed, single step skill.

Although this procedural stage is very similar to Shiffrin and Schneider's automatic processing, there are important differences. It is conceptually more complex because Anderson offers more detail concerning how proceduralised knowledge may develop. As knowledge moves from a declarative to a procedural form it is changed by two processes that Anderson calls *compilation* and *tuning*. In the process of compilation separate bits of information (e.g. digits) that are required in sequence (e.g. a phone number) in the performance of some action (e.g. dialling that number) are chunked together, requiring less effort to execute. (Which explains why we speak and write phone numbers in groups of three or four digits rather than as strings of individual digits). These easier-to-execute chunks, or *compositions*, are then proceduralised, which means that they can be retrieved from long-term memory without the need to activate conscious declarative knowledge. (Which explains why we can dial a well-known phone number without any conscious rehearsal of the digits.) In the process of tuning, procedural knowledge undergoes further transformations that enables it to be generalised, discriminated or strengthened.

While this has no application to the learning of simple skills (like memorising phone numbers) it has obvious relevance to the learning of complex skills such as language. The acquisition of a first or second language requires the learner constantly to review previously learned knowledge in order to incorporate newly learned knowledge. It also requires the learner to be able to replace proceduralised knowledge that turns out to be wrong. In this respect, ACT* is a more satisfying model for language learning than Shiffrin and Schneider's in which automatized processes are basically unchanging, speeded-up versions of controlled processes. ACT* can also explain why language learning is not usually completely successful. According to ACT*, the extent to which a skill continues to be finely-tuned over time is a function of the amount of mental effort this takes balanced by the amount of progress it achieves. This means that the *tuning* of a particular skill will cease when progress in accuracy begins to be outweighed by the effort it demands. This limit to the restructuring or reorganisation of knowledge has important echoes in second language acquisition, where it is termed *fossilisation*. For some L2 learners this limit is reached early, others late, but almost all eventually cease learning

before they have achieved perfect knowledge. This could well be because the effort involved in fine-tuning their existing knowledge is greater than the benefits any extra knowledge might bring. For learners of English L2, for example, a better understanding of the complexities of the article system brings little communicative advantage and might therefore seem not to be worth the trouble. But why some learners nevertheless master conceptually complex but communicatively redundant L2 forms cannot be accounted for by ACT*.

Although there remains disagreement as to the exact nature of the way the brain learns, stores and retrieves information, the notion of the brain as a limited-capacity information processor is widely accepted as correct. To combat the difficulties caused by its inability to pay conscious attention to more than one thing at a time, the brain has to develop automatic responses (or procedures) which, while not always capacity-free, nevertheless require so little capacity that they allow simultaneous conscious attention to be given to controlled processes. Only in this way can the execution of complex skills be properly accounted for. Complex skills are characterised by a hierarchy of component sub-skills, the higher orders of which are dependent on the successful, automatic execution of lower order components. In the case of speaking, the lower order skills include the ordering and articulation of selected phonemes, the retrieval of selected vocabulary, the execution of appropriate syntactic and morphological rules, and the meeting of any necessary pragmatic conventions. (Levelt 1978). For the higher order skill of expressing a particular intention to be possible, all these sub-skills must be automatised. In other words, for us to be able to think about what we are saying, we must be freed from the necessity of thinking about how phonemes, words and phrases are constructed. In the following sections we shall consider the extent to which it is possible for us fully to automatise these lower order speaking skills, or the extent to which they may continue to demand some cognitive capacity.

2.3 The role of automatization in first language performance

ACT* has been extensively tested and found to be a robust model for the learning of algebra, geometry and computer programming. It is intuitively appealing (and supported by research evidence) that learning how to solve, say, algebra problems is initially the painstaking and uncertain application of mathematical principles, and ultimately the rapid and confident application of the same principles. But Anderson's (1983) attempt to apply ACT to first language learning is more problematic. A production such as:

IF the goal is to generate the plural of man,
THEN say 'MEN'.

seems straightforward enough, but the production sequence necessary to provide the correct verb form in 'the dog chases the cats' should make most linguists uneasy.

IF the goal is to indicate the relation in
(LVobject 1 chase LVobject 2)*
and LVobject1 is dog
and LVobject1 is singular
and LVobject2 is cat
and LVobject2 is plural
THEN say 'chases'

**LV = linguistic variable* (Anderson 1982:391)

According to Anderson, *all* knowledge starts out as declarative, i.e. explicit and capable of verbal mediation. It then becomes proceduralised and may, as a result, be lost as declarative knowledge. (It is possible, for example, for us to be able to dial a very frequently used phone number but have trouble saying what the number is, or to be able to play a piece of music without being able to write down a single bar.) In the case of language however, to assert that we as children knew explicitly the rules of our language but have since lost conscious access to them, is patently absurd. Anderson has since withdrawn from this extreme position (Anderson and Finchum 1994).

It is more profitable to look for evidence of automaticity in the subskills of language production, such as the articulation of phonemes, and word retrieval. For young children beginning to learn language, these subskills require significant investments of processing capacity. In order for children to progress beyond the one-word-at-a-time stage it seems clear that some automatisation, (or proceduralisation) takes place that allows the child to allocate processing resources to the higher and less predictable skill of syntactic processing. There is some good evidence for this. Scollon (1976) reports that his subject was less successful at articulating the same words in longer utterances than in shorter ones, indicating that the effort required to construct a longer utterance took attention away from the effort required to pronounce it. Bloom, Miller and Hood (1975, reported in Bock 1982) found that the children in their study produced longer utterances when using vocabulary that was very familiar to them than they did when using vocabulary that was relatively new, indicating that using unfamiliar vocabulary diverted their attention from constructing the utterance. As phonological and lexical processing become automatised in the child, they require less and less capacity and no longer impair the performance of other language skills.[2] There is evidence that children learning language have a propensity to develop automatic responses to linguistic information because they assume it is *consistently mapped*. They overgeneralise morphological patterns such as the past tense *-ed* and plural *-s* to produce such forms as *putted* and *foots*, expecting patterns in morphology to be regular and predictable. The fact that all children go on to reanalyse these incorrect rules and transform them into the correct ones is good evidence for Anderson's *tuning* in the learning of a skill.

Phonology and lexis lend themselves comfortably to consistent mapping and thus to automatic processing. The same combination of muscle movements in the tongue, lips and glottis produce a particular sound, the same meaning is, usually, conveyed each time by a particular word. The Stroop test (Posner and Syder 1975) demonstrates that there is indeed automaticity in lexical retrieval. Subjects were shown the names of colours written

---

[2] For adults, phonological patterning and lexical retrieval seem effortless, unless we are asked to repeat a tongue-twister at speed, or fail to find the word we are looking for and have to wrestle painfully with it on

in ink of another colour. They were often unable to suppress saying the word written on the card, even when instructed to say the colour of the ink. Favreau (1981) found a difference in bilingual subjects in their speed of word recognition in their first language compared to the speed of word recognition in their second language, indicating that they were not processing words from their second language as efficiently (automatically) as words from their first language. This was not the case for balanced bilinguals (i.e. those with two first languages) who showed an equal facility for recognising words from both of their languages. Balanced bilinguals also showed no difference in reading speed in their two languages, whereas dominant bilinguals (i.e. those with one first language) were slower in reading their second language, again demonstrating that the processing of the first language was faster and more successfully automatised.

There is evidence that the decoding of syntax is highly automatised. As long ago as 1967 Sachs demonstrated that native speakers cannot easily recall the syntactic form of an utterance as well as they can recall its semantic content. Much of the syntactic structure of the language we hear every day decays from our memory almost as soon as we hear it, even when we are determined to remember it, as anyone who has tried to learn a poem by heart can attest. Hatch, Polin and Part (1970) asked native speakers to cross out all instances of the letter 'e' in a text, and found that they were significantly more likely to miss the 'e's in function words than in content words, suggesting that they were using automatic processing for the syntactic elements of a sentence and more controlled processing for the semantic elements. As for *en*coding language in speech, Bock (1982) believes that some syntactic patterns exhibit characteristics that could be accounted for by automatic processing: the basic phrase types common to all languages (noun phrase, verb phrase and prepositional phrase); in English, the subject-verb-object word order and the overwhelming preference for active over passive verbs. However, this degree of regularity accounts for only a very small proportion of the syntax any speaker regularly produces. There are many choices to be made in the syntactic form of an utterance which allow a speaker to tailor utterances to the communicative demands of the situation.

---

the tip-of-the-tongue. This takes all our attention and can be very tiring.

The window is broken.

It looks like the window has been broken.

Does that window look broken to you?

I think I might have broken the window.

I broke the window.

All of these sentences are conceptually very similar, but syntactically (and pragmatically) very different. They cannot be said to have been automatically processed in response to an accident with a window, but purposefully produced by the speaker after due consideration of what would be most appropriate. In other words, their production certainly involved some degree of *controlled* processing. This means that for much syntax there is a cost in terms of processing capacity, not just because syntax is difficult (and it certainly can be) but because its production cannot be fully automatised. This cost is unlikely to show up in adults as a falling off of performance in phonology or lexical retrieval as both are so well automatised, but it is likely to show up in that other area of controlled processing in language production: content planning. A variety of studies (e.g. Goldman-Eisler 1968, Deese 1980) have shown that speakers who are having to cope with an unfamilar or difficult subject matter become dysfluent as they attempt to attend to the form of their speech and its content at the same time, again suggesting that our controlled processing resources are limited and cannot be divided without loss. Bock (1982) suggests that when linguistic processing resources are stretched (because of demanding subject matter and/or great time pressure) a 'default mode' of syntactic processing may be invoked. She imagines these as a set of syntactic patterns or rules used by speakers when they have insufficient attentional resouces to deal with the form and the content of what they are saying. Bock does not speculate what these 'default' syntactic patterns might be, but her idea is close to those of Ochs (1979) and Givon (1979) which are discussed below in 2.4.

Another perspective is offered by Pawley and Syder (1983). They point out that a very important feature of native speaker language is that it is not only syntactically well-

formed, but it also *sounds natural*. The grammar of any language can generate an infinite number of permissable utterances but most of them would not be accepted by native speakers as normal. For example, the natural (and attested) utterance,

(1). I'm so glad you could bring Harry!

can be recast as:

(2). That Harry could be brought by you makes me so glad.

(3). That you could bring Harry gladdens me so.

(Pawley and Syder 1983:195)

but these are legitimate only as far as the rules of sentence construction are concerned. Examples (2) and (3) would be noted by any native speaker of English as meaningful, but very odd and unnatural. Pawley and Syder maintain that knowing a language is more than knowing its rules of construction, it is also knowing what constructions are idiomatic and natural. This, far from being an added burden to the language learning process, is the beneficial result of regular form-meaning pairings being recognised and stored in memory. A child learning his first language will notice regularly occurring lexical and syntactic patterns in the language spoken around him and will learn them. As a adult native speaker he will know many thousands of such patterns and will draw on them regularly when encoding and decoding language. Pawley and Syder claim that these patterns are either completely lexicalised phrases:

'I told you so.'

'Never mind'

What's the matter'

or lexicalised sentence stems:

Who (the EXPLETIVE) do- PRES NP think PRO be-PRES

→ who the hell do you think you are.

→ who the devil does she think she is.

→ who the blazes do they think they are.

(after Pawley and Syder 1983:211)

These stored sequences have the advantage of being familiar to the listener and to the speaker and can account, in part, for the fluency with which spoken language is delivered and understood. Equipped with thousands of these ready-made or partially-made chunks of language, the speaker is able to encode his meaning more rapidly than if he had to construct all utterances from scratch, and the listener, equipped with the same knowledge of these chunks is able to anticipate the syntactic and lexical direction of an utterance before it has been completed.

The notion of lexicalised chunks and lexicalised stems accounting for the ease and speed with which spoken language is delivered is paralleled in Shiffrin and Schneider's model of learning by the notion of connections between nodes in the long-term memory store being strengthened by repeated exposure and practice. Lexicalised language is thus automatised language, *not in the sense of sequences of words being triggered automatically and* irrepressibly without the conscious will of the speaker, but in the sense of familiar sequences of words being processed rapidly and effortlessly as unanalysed wholes (in the case of chunks), or unanalysed frameworks (in the case of stems), allowing the speaker to attend to how these can be syntactically completed and fitted together.

> In the store of familiar collocations there are expressions for a wide range of familiar concepts and speech acts, and the speaker is able to retrieve these as wholes or as automatic chains from the long-term memory; by doing this he minimises the amount of clause-internal encoding work to be done and frees himself to attend to other tasks in talk-exchange, including planning of larger units of discourse. (Pawley and Syder, 1983: 192)

Ellis (1996) also argues that much of first and second language learning is the memorisation of chunks or sequences of language, and that fluency is the result of access to these sequences becoming more automatic. He further argues that the possession of a long-term memory store of word sequences enables a language learner to acquire a knowledge of grammar by analysing patterns in the sequences and seeing the regularities that govern them. In this way, the Shiffrin and Schneider model of learning is accommodated into Anderson's ACT* model, with the long-term store of lexicalised chunks gained through repeated experience providing the data-base from which declarative knowledge of grammatical patterns is derived and, ultimately, tuned and proceduralised.[3] This interesting angle on the role of pattern memorisation and analysis in language learning will be discussed at greater length in chapter eight below.

2.4. Processing constraints in first language performance.

At about the same time Shiffrin and Schneider were demonstrating in their research in psychology that human beings were limited-capacity information processors, researchers in linguistics were independently formulating very similar hypotheses. For these linguists (e.g. Kroll 1977, Givon 1979, Ochs 1979) the main interest was not in the mechanics of information processing, but in how the constraints of limited attention are manifested in language production. All conclude that the more complex syntax is, the more attentional capacity it demands.

The widest perspective was taken by Givon (1979). From a detailed consideration of common patterns of language change, the evolution of syntactic creoles from asyntactic pidgins, and the features of both first and second language acquisition, he concludes that grammar itself is an automatic processing strategy. He argues that there is an amazing uniformity across languages in the evolution of tense and aspect auxiliaries out of loose

---

[3] Declarative knowledge of language structure, especially in the case of L1 acquisition, could not mean conscious, explicit knowledge of language structure. It can only refer to 'knowledge that' before it becomes proceduralised into 'knowledge how'.

paratactic constructions, and that these markers may become bound morphemes of the verb. He gives examples of the most common changes:

want	→FUTURE
go	→FUTURE
come	→PERFECTIVE → PAST
finish	→PERFECTIVE → PAST
have	→PERFECTIVE → PAST
be	→PROGRESSIVE/HABITUAL → FUTURE
start	→FUTURE
can	→HABITUAL/POSSIBLE/PERMISSIBLE
lack/fail/refuse	→NEGATION
done	→PERFECTIVE → PAST

(after Givon 1979: 96)

He shows that in many languages there is evidence that the major, tightly-bound subordinated constructions arose from conjoined constructions, i.e. that language constantly takes discourse structures and condenses these into syntax. In English, such a process would include the use of the verb *will* to express future time. In Tok Pisin, an English-derived language in Papua New Guinea, *bin* (been) is used as a past tense marker for the verb e.g. *'asde mi bin go = I went yesterday'*. (This process, which would otherwise lead to more and more complex morphology, is offset by phonological attrition whereby unstressed morphemes tend to be lost over time, as in the noun case endings of Anglo-Saxon.) In the terms used by psychologists (which Givon himself did not use) one can say that the controlled process of a paratactic construction becomes the automatic process of a syntactic construction, requiring thereby less processing capacity.

Givon sees a similar syntacticization process at work in the development of creole out of pidgin. Pidgin languages have no native speakers and are used as a means of communication between peoples who otherwise have no common language. They exhibit

such wide variation in their structure that they can be said to have no stable syntax at all. Word order can be inconsistent and grammatical morphemes very few. (Todd 1984) As a result, utterances are cumbersome, having to be made up of simple clauses, most commonly of the *topic-comment* type with a one-to-one ratio of nouns and verbs, strung together in loose co-ordination. Creoles, by contrast, are the result of children receiving pidgin as their mother tongue. Within five to seven years they transform this limited input into a fully-fledged language i.e. a system with extensive syntactic and morphological patterns that is able to express complex propositions quickly and efficiently. Givon describes this transformation as the development of *automated coding processes.* (1977:111). Clearly, this is not the *automatisation* described by Shiffrin and Schneider, but something more akin to the *restructuring* described by Cheng. It is interesting that this shift from cumbersome co-ordination to complex syntax is achieved only by children and not by adult speakers for whom a pidgin may speed up with practice (i.e. be automatised) but will not change (i.e. be restructured) into a more efficient language system.[4]

Givon compares pidgins to the language of young children and also to *foreigner talk*, the style adults adopt when talking to someone who is not fluent in their language. All of these modes of speech display common features: a lack of grammatical morphology, a reliance on co-ordination rather than subordination, signalling new information through higher intonation, and a noun to verb ratio of almost one-to-one. For Givon this is the *pragmatic mode*, what he calls 'the bottom line register shared by all humans'. The communicative situation that prevails for pidgins, foreigner talk and child language is remarkably the same: there is some communicative stress (i.e. no common, fully developed language through which to communicate), no common pragmatic background (i.e. a shared culture or understanding of social conventions), but there is an immediately obvious context (i.e. the topic and the task are in the here-and-now). As the child matures he gains linguistic and social skills. He acquires more and more of the pragmatic conventions of his culture and he is able to deal with topics and tasks that are not rooted

---

[4] That the ability to syntacticise is possessed only by pre-pubescent children is well attested in studies of creolisation and sign language.(see Pinker 1994)

in the immediate context. His language moves beyond the pragmatic mode into the *syntactic mode*, with elaborate use of grammatical morphology, a higher ratio of nouns to verbs, a tighter subordinate structure and less reliance on intonation to signal new information. The two modes are summarised in the table below.

Pragmatic Mode	Syntactic Mode
topic-comment structure.	subject-predicate structure.
loose conjunction.	tight subordination.
slow rate of delivery (under several intonational contours).	fast rate of delivery under one intonational contour).
old information goes first, new information follows.	word order is used to signal semantic case functions.
roughly one-to-one ratio of verbs to nouns, with verbs being semantically simple.	larger ratio of nouns over verbs, with verbs being semantically complex.
no use of grammatical morphology.	elaborate use of grammatical morphology.

(adapted from Givon 1979:109)

According to Givon, the syntactic mode does not supplant the pragmatic mode. The two co-exist as opposite ends of a continuum that covers a whole range of styles from the extremely syntactic (as might be used in formal writing) to the extremely pragmatic (as might be used in trying to communicate with a foreigner under stressful conditions). The key to the selection of a pragmatic or syntactic mode is, for Givon, the question of communicative stress and time pressure. Informal conversations under relaxed conditions where there is face-to-face monitoring can afford to be less tightly organised, less planned, more pragmatic in mode. Spoken language such as a public speech or a statement to the press is unlikely to be delivered extemporaneously in a pragmatic mode because in such a situation language needs to be more formal, more carefully organised, more syntactic in mode. Written language has to communicate without the assistance of face-to-face monitoring, intonation and gesture, but is produced without the time pressure that

accompanies spontaneous speech. It can be cast and re-cast in complex syntax until the writer is satisfied with it. Equally, an extreme emergency in which the immediate communication of information is of the utmost importance *requires* the use of the pragmatic mode because there is no time and no attention for organised, planned, syntax.

Givon's main point is that grammar itself is a processing strategy. It speeds up the cumbersome process of producing loosely co-ordinated topic-comment structures, and allows us to communicate complex information efficiently and quickly. But it is important to note Givon does not say that the gains in efficiency and speed are without cost. In the psychology terms we have used above, the syntactic mode is *not* an automatised, capacity-free version of the controlled capacity-demanding pragmatic mode. Givon sees the syntactic mode as more capacity-demanding than the pragmatic mode and this is why, when time is very short, or the situation very pressured, we do not use it. Although the use of the syntactic mode increases the processing burden, it allows us to express more complex meanings with greater concision and precision. Syntax is the *restructuring* of pragmatics, with gains in range and speed, and losses in attentional capacity.

Ochs (1979) had independently arrived at the same conclusion, although by a rather different route. Her research focussed on a comparison of formal, planned language with more informal, unplanned language in order to discover what the syntactic differences might be. She was concerned not only with the processing demands of simple and complex syntax, but also with the way a speaker's attention to the syntactic coding of a message must compete with his attention to its content, and its context.

For Ochs, it is obvious that a speaker may face conceptual and/or situational demands which reduce the amount of attention that he or she is able to devote to all dimensions of the message form. In spontaneous conversation, for example, situational demands would include monitoring the speech and body language of the other participant(s), anticipating 'transition relevance points' (Sacks, Schegloff and Jefferson 1974) for when it is acceptable to begin speaking, and observing whatever other pragmatic conventions apply

to topic holding and topic changing. Conceptual demands arise when the subject matter to be encoded in speech is so unfamiliar and/or complex that it requires a lot of concentration on the part of the speaker. These two kinds of demand may well compete with each other. A speaker may be so concerned with the demands of the situation, e.g. a conversation where he is very anxious to observe social conventions, that he is unable to give proper attention to the subject matter. Equally a speaker may be so focussed on working out an idea that he may neglect the normal turn-taking conventions by, say, lapsing into silence, butting in on another speaker or talking for far too long. But for Ochs, the most interesting result of heavy situational or conceptual demands is the impact these have on syntax.

Ochs theorises that complex syntax requires *planning* before it can be executed, and that consequently the brain must allocate extra attentional resources to it. When this is lacking due to situational or conceptual demands, only simpler syntax can be produced because this does not require planning. She further theorises that this simple syntax consists of the morphosyntactic and discourse skills that are learned in the first three to four years of life. These skills are not replaced by more complex skills as the child matures linguistically, they remain throughout adulthood as a resource to be called upon when needed. Thus the discourse of adults which is produced without any forethought or organisational preparation, (as you would probably find in a spontaneous conversation) will share many of the morphosyntactic and discourse features of the speech of young children. Equally the language of adults which has been thought out and organised beforehand,(as you would get in a speech, or in a piece of writing) will be morphosyntactically and discoursally mature.

She tested this hypothesis by examining a data base made up (in a rather ad hoc manner) of child-child recordings, child-adult recordings, and spontaneous conversations between adults, all collected over a period of time either by Ochs herself or by colleagues. In addition Ochs made recordings of personal narratives delivered under two conditions by the same adult speakers. In the first condition, the speakers had to relate an incident in

their lives in which they had been in danger to an audience of fellow students. This was done without any preparation or rehearsal and was deemed to be relatively unplanned. Each speaker had then to write the same incident down and hand it in within two days. This was deemed to be relatively planned. All this assembled data was analysed for features that would distinguish relatively unplanned discourse (the adults' conversations, the spoken narratives) from relatively planned discourse (the written narratives), and that would be shared by the discourse of young children.

One can say at once that Och's way of gathering and analysing data is so haphazard that any conclusions she draws from her study are at best rather tentative. As we have noted, the data was mostly an ad hoc pooling of recordings made at various times with no effort to control intervening variables. She does not define many of the terms she uses (e.g.'social act' 'repetition'), she does not justify their psycholinguistic significance, and worse, she does not quantify their occurrence, preferring always to say that something appears 'with greater frequency' in one condition or the other. Even more seriously, she has used writing as the only example of planned discourse, and contrasted this with many spoken examples of unplanned discourse. Consequently, she is not comparing like with like, and it would hardly be surprising to find that writing is different from speaking.

Ochs' conclusions therefore must be viewed with a certain amount of circumspection. But although we can say that her study has not satisfactorily demonstrated anything at all, we can nevertheless treat her observations as worthy of consideration and further investigation. From her analyses she identifies four features that distinguish relatively planned from relatively unplanned discourse. Feature four, which is by far the least well explained, is to do with the relative compactness of planned discourse compared to the relative wordiness of unplanned discourse. The difference is not accounted for by Ochs as resulting from competition for attentional resources, and so need not concern us here.

1. *In relatively unplanned discourse more than in planned discourse, speakers rely on the immediate context to express propositions.*

2. *In relatively unplanned discourse more than in relatively planned discourse, speakers rely on morphosyntactic structures acquired in the early stages of language development. Relatively planned discourse makes greater use of morphosyntactic structures that are relatively late to emerge in language.*

3. *In relatively unplanned discourse more than in relatively planned discourse, speakers tend to repeat and replace lexical items in the expression of a proposition.*

(after Ochs 1979)

Features one and two are concerned with the allocation of attentional resources for syntactic production. When attention to language form is limited by competing situational or conceptual demands, the immediate context can be used instead of a syntactic link, e.g. *'I don't like that house. It looks strange'* as opposed to '*I don't like that house because it looks strange.'* Also when attention is thus limited, more difficult syntactic structures such as the passive voice or complex past tenses are avoided in favour of the simpler active voice and present tense, e.g. *'she's talking to a friend'* as opposed to *'the friend to whom she was talking'*. Feature three is concerned with the allocation of attentional resources to lexical retrieval. When attention is limited there is more likelihood that the speaker will repeat words because he isn't sure of what to say next, or replace words because the initial selection was inappropriate or inaccurate. *'There were these people talking, and this woman-- **lady** was describing something....'* (Examples adapted from Ochs 1979.) Ochs believes that features one to three are also characteristic of the language of young children, and their development of control over more mature language can be linked 'to an increased capacity of the child to attend to both the form and the content of the propositions they express.' (p.73)

Kroll (1977) made a quantitative analysis of the seven spoken and written narratives used by Ochs (1979) in order to determine if there was a difference in the incidence of co-

ordination and subordination between the two conditions. She analysed the data into 'idea units'. These are defined as a phrase, clause or sentence that represents a chunk of information viewed by the speaker or writer as an inviolable unit. How these units are strung together should reflect the amount of attention the speaker or writer paid to the form (rather than the content) of the discourse. Co-ordination, which for Kroll is signalled either asyndectically, or by the use of conjunctions (e.g. *and, but, so*), is assumed to need less attentional capacity than subordination, which is signalled by the use of subordinators (e.g. verb + ing, preposition + ing, relative pronouns, because, etc.). The results show that in the spoken narratives there was a strong preference for the use of *'and'* to co-ordinate idea units, (on average, 36% of units were co-ordinated in this way) and an equally strong preference not to link units at all, (on average 38% of units were not linked in any way to the preceding unit). Subordinators were rare (linking on average only 7% of units) and co-ordinators other than *'and'* were rarer still (linking on average 4% of units). For the written versions the results are very different. While there is still a significant number of units that are not linked to the preceding unit (40%), the use of subordinators rises to an average of 19%, and co-ordinators fall to an average of 25 %.

Kroll concludes that the spoken narratives relied so heavily on the use of *'and'* to co-ordinate ideas because the time and situation pressure allowed no attention to be given to the major manipulations of syntax and morphology required for combining ideas through subordination. The written versions were produced under conditions that gave the writers ample time to attend to these manipulations. However Kroll's study is open to the same criticism as Och's: she is comparing speech with writing. It is quite obvious that writing gives time for reflection upon syntactic structures, especially when the writing is done over a period of days and without any time pressure. This does not mean that extemporaneous speaking gives insufficient time for reflection upon syntactic structures and therefore must be delivered in simple syntax. The fact that the subjects avoided subordinated structures in their spoken narratives could be explained by the way they viewed their two tasks. They had first to tell their classmates about a time when they had been in mortal danger. Many people would agree that short, co-ordinated structures

accompanied by appropriate intonation, pausing and gestures is a good way to describe the drama of a situation to an audience. Complex subordinated structures in such a setting might well seem inappropriate. The same students, writing down their story, did not have an immediate audience and no longer had surprise or suspense to convey because the story had already been told. The task was therefore very different, and the students almost certainly had notions about what makes a 'good' piece of writing, i.e. a more formal register involving a greater attention to detail and a more complex structure. Moreover, as each of the stories was about a life-threatening event it is inconceivable that the students had not already recounted the story many times before. In this case the narratives are not examples of unplanned discourse, and the fact that the students still chose to speak them in relatively simple co-ordinated structures shows that the choice was deliberate, not forced upon them by the pressures of the situation. If Julius Caesar ever did exclaim, '*I came, I saw, I conquered*' it was not necessarily because he had insufficient attentional resources to say instead, '*After I had come and seen, I conquered.*' In spite of having all the time in the world to think about his written version, he rightly chose the former for its dramatic impact.[5]

Neither Kroll nor Ochs has demonstrated that there are syntactic differences between relatively planned and relatively unplanned discourse because neither of them set about gathering and analysing data in a principled way. But they have raised interesting research questions. Data could be gathered from subjects speaking extemporaneously on some topic, and this could be compared with data from a similar group of subjects who had prepared to speak on the same topic. It would then be possible to look for syntactic and lexical differences that could properly be accounted for by the absence or presence of planning time. Unfortunately, since the work of Ochs and Kroll in the late 1970s, only linguists interested in second language acquisition have pursued lines of research in this area. We shall look at this work in Chapter Three, leaving the research into planned and

---

[5] That is to say, *veni, vidi, vici*, is a simpler and more dramatic way of expressing *cum venissem et vidissem, vici.*

unplanned native speaker discourse as it remains today: largely speculative and with only a small and shaky empirical base.

Chapter Three

Information processing and second language learning.

It is uncontroversial to observe that second language use is characterised by slow effortful performance being gradually replaced by faster and less effortful performance. On the surface this would seem to indicate that, in common with the learning of other skills (cognitive and motor) some automaticisation (or proceduralisation) must be taking place. Exactly how this process operates in second language learning, however, is controversial, and has been the focus of a great deal of argument and research. The debate has come to centre on whether knowledge of the target language structure is implicit or explicit (or a mixture of both), and on whether learning takes place with or without the conscious awareness of the learner. Much of the controversy has to do with the precise definition of terms. What exactly do we mean by 'conscious awareness' in language learning? What precisely do we mean by 'learning' anyway, and is it different from 'acquisition'? What is meant by 'implicit' or 'explicit' knowledge of linguistic structures, and by the 'implicit' and 'explicit' learning (or acquisition) of them? In this chapter we shall consider the competing theories for the roles of implicit and explicit knowledge in second language performance, and the different descriptions of the nature of consciousness. Although we shall consider these two issues separately, they are two sides of the same coin, the first concerned with the product of second language acquisition, and the second concerned more with its process. Both ultimately involve the further issue of the way second languages should be taught, which will be the focus of chapters four and five.

3.1 Implicit and explicit knowledge of language.

Native speakers are able to demonstrate their knowledge of the grammar of their language by their fluent and accurate production of it. They can also demonstrate this knowledge by correctly judging examples of the language as grammatical or ungrammatical. However, native speakers who have not applied themselves to the study of the grammar of their language are most unlikely to be able to explain just why their language requires words to

be in a certain form or order, or why other word forms or word orders are unacceptable. An example which never fails to baffle native speakers of English is as follows:

1. After Mary came home, she had a shower.
2. After she came home, Mary had a shower.
3. She had a shower after Mary came home.
4. Mary had a shower after she came home.

These four sentences contain the same eight words, but each has them in a slightly different order. Three of the sentences might refer to just one woman, Mary, or might refer to two women, Mary and someone else, but the remaining sentence cannot refer to one woman and must refer to two. Not many native speakers have any trouble in identifying (3) as the only one which must refer to two women, and none at all would ever produce (3) when intending to refer to Mary alone. But that a native speaker untutored in English grammar could ever articulate the rule that a pronoun never refers to the noun in a following dependent clause is almost inconceivable[6]

Rules such as these exist in their thousands in every language and are, in an implicit sense, known perfectly by their native speakers. However, making these rules explicit, i.e. able to be discussed and reflected upon, requires the invention of a special metalanguage to describe them and is the laborious work of specialists. The vast majority of human languages have kept the intricacies of their structure locked away in implicit knowledge.

Children learn their first language(s) by gaining implicit knowledge of the grammar. Their caretakers do not know the rules explicitly, cannot explain them and therefore cannot teach them. There is plenty of evidence that those parents (usually linguistic researchers) who have tried explicit language instruction on small children were completely wasting their time (see Braine 1971 for some examples). Children learning their L1 are simply not

---

[6] Almost inconceivable because this elusive rule defied formulation until Langacker managed it in 1969. (Trask, 1995)

receptive to explicit information about the language. There is now a very strong case (see Pinker 1994 for a full review) that humans have a genetically endowed 'language acquisition device' (Chomsky 1965) that enables them (in the early years of life only, alas) to get full implicit knowledge of a language extremely rapidly and, moreover, without the slightest awareness that they are doing so. Implicit knowledge of L1 grammar is an unsought and unconscious by-product of children's conscious efforts to do something else, namely, communicate with those around them.

Adults learning a second language are exceedingly unlikely to be unaware that this is what they are doing. For those learning in a classroom, or from a teach-yourself book, the structure of the target language is often presented quite explicitly, with rules that the learner is expected to commit to memory.

> The present perfect tense is formed with the present tense of **to have** + the past participle: *I have worked,* etc.....
>
> The present perfect is used for past actions whose time is not given and not definite.
>
> A It is used for recent actions when the time is not mentioned:
> *I have read the instructions but I don't understand them.*
>
> B It can also be used for actions which occur further back in the past, provided the connection with the present is still maintained, that is, the action could be repeated in the present.
> *John Smith has written a number of short stories* implies that he is alive and could write more stories, whereas if John Smith were dead we would say, *John Smith wrote a number of short stories.*
>
> Etc.
>
> (from Thompson and Martinet 1980: 51-152)

As a result of explicit teaching, the learner is able to talk about the target language, confidently describing rules and their exceptions in a way that untutored native speakers cannot match. Such explicit knowledge of linguistic rules is supposed to guide the learner's attempts to understand or produce target forms accurately. In Anderson's ACT* model this explicit knowledge is declarative knowledge, capacity-demanding and characterised by slow, effortful performance. As this knowledge is proceduralised, performance speeds up and requires less processing capacity. Whether or not declarative knowledge is ultimately lost from conscious access as it is proceduralised (i.e. becomes like the inaccessible implicit knowledge of native speakers) is a moot point. Anderson certainly believes this to be possible. What is important is that proceduralised knowledge, accessible or not, is like implicit knowledge of a first language: it is the *only* basis for fluent and accurate language production. Some theorists (e.g. Krashen 1981, 1982) maintain that implicit knowledge of structure (i.e. that which makes fluent and accurate language production possible) never arises from an earlier explicit stage, and that explicit knowledge of language structures is not only entirely unnecessary to the learning process, it is even harmful to it. It is to the details of this debate, and its relevance to second language acquisition, that we now turn.

3.2 The interface between explicit and implicit knowledge.

In Ellis, R's (1993) analysis, there are three schools of thought on this topic: those who consider that explicit knowledge cannot be turned into implicit knowledge (the 'no-interface' position), those who think that explicit knowledge can assist in the development of implicit knowledge (the 'weak interface' position), and those who think that explicit knowledge is converted into implicit knowledge through practice (the 'strong interface' position).

3.2.1 The 'strong interface position'

This is based upon the research evidence from experimental psychology, reviewed in chapter two above, that cognitive skills are acquired (automatised/ proceduralised) through practice. The information-processing models of both Shiffrin and Schneider (1977) and Anderson (1982) describe a process in which slow and deliberate performance is gradually replaced by performance that is considerably faster and more efficient. Both models are based on extensive research evidence from search-and-detect tasks and solving geometry problems rather than research into language learning, but both have been generalised to include language. McLaughlin et al. (1987:153) for example are quite specific on this matter.

> .....in our conception of the language learning process, repeated performance of the components of the task through controlled processing leads to the availability of automatised routines.

There are two crucial questions, however, for anyone taking a 'strong interface' line: 1) are first language skills uniquely structured and therefore impossible to fit into models of learning based on other cognitive skills, and 2) are second languages learned in the same way as first languages? On the first question, Anderson has shown himself very unwilling to entertain the idea that first language acquisition may not fit ACT*. Although he has moved from his earlier (1983) insistence that first language learning is a process of proceduralising declarative knowledge through practice, his assertion that, 'Children ...appear not to be much helped by direct instruction but rather learn language more implicitly' (1990:325) would be viewed as unjustifiably grudging by any researcher into first language acquisition, all of whom would agree that the implicit nature of children's language learning, and the futility of direct instruction (not to mention its absence in real life) are axiomatic. Anderson's further assertion (1990:356) that, 'it is a fair summary to say that the jury is still out' on the issue of whether language is different from all other

cognitive systems is likewise out of tune with current thinking in applied linguistics. For most linguists, the jury came back some time ago.[7]

The second question, as to whether first and second languages are learned differently, is more interesting. If Lenneberg's (1967) critical period hypothesis is right, language learning after puberty has to be accomplished without the help of the innate language acquisition device because this will have ceased to function. The effortless speed with which young children come to master their first language is quite obviously not repeated in adults learning a second language. That the two processes are therefore fundamentally different is a very plausible conclusion. As we shall see below in chapter four, most traditional (and much contemporary) language teaching assumes that second language learning is essentially the same as learning any other cognitive skill and is achieved through the gradual automatisation of explicit rules. Skilled and fluent performance of the target language is supposed to come about because repeated practice allows the application of explicit rules to become quicker and less effortful. According to Sharwood Smith (1981: 66),

> ...it is quite clear and uncontroversial to say that most spontaneous performance is attained by dint of practice. In the course of actually performing in the target language, the learner gains the necessary control over its structures such that he or she can use them quickly without reflection.

Failure to automatise the rules is viewed as the result of poor understanding (requiring more explicit explanation) and/or insufficient practice (requiring more drilling or other performance exercises).

---

[7] Anderson and Finchum (1994) retreat from this earlier position, but Anderson (1995:378) is adamant that language should not be treated as qualitatively different from any other skill.

3.2.2 The 'no-interface' position

According to this, second language acquisition is *not* fundamentally different from first language acquisition, and the explicit teaching of grammatical structures is therefore neither necessary nor desirable. The best known advocate of the 'no-interface' position is Krashen (1981, 1982, 1985). He makes a fundamental distinction between 'learning' a language (which will result only in explicit knowledge of its structure) and 'acquiring' a language (which will result only in implicit knowledge of its structure).[8] Explicit knowledge may be perfectly well learned, but it will not result in implicit knowledge and thus will not aid fluent and accurate L2 production. If a learner 'learns' a target language rule at some stage (i.e. explicitly) and then at some later stage shows that he can use this rule in fast and accurate performance, it is not because the explicit knowledge has been proceduralised, but because implicit knowledge of the rule has been 'acquired'. This 'acquisition' of the rule has come about through exposure to sufficient comprehensible input, i.e. samples of the target language which the learner was able to understand and which contained examples of the rule in action. Krashen maintains that research showing learners can use rules of the target language without being able to explain what these rule are is solid empirical evidence for implicit knowledge having no source in explicit knowledge. Many critics (notably Gregg 1984 and McLaughlin 1987) have pointed out that the inability to give a verbal account for a rule is no evidence whatsoever that its structure was never consciously attended to, wrestled with or despaired about. In fact, Krashen's critics have shown that the separation of acquired (implicit) knowledge from learned (explicit) knowledge is empirically unfalsifiable. As we can never isolate implicitly acquired rules from explicitly acquired rules, we have no way of knowing if there are indeed two separate and unconnected systems. Nevertheless, the 'no-interface' position still has its advocates in the language teaching profession.

3.2.3 The 'weak interface' position

The middle way on this issue is that there is a weak interface between implicit and explicit knowledge (Seliger 1979, Ellis, 1993a). This assumes that explicit knowledge of language

---

[8] The distinction in the meaning of these terms is not held in this thesis.

structures can become implicit knowledge, but only under certain circumstances. Knowledge which is gained through explicit teaching can help learners to become aware of certain features in the target language, (which they might otherwise have failed to see) and thus facilitates their subsequent acquisition as implicit knowledge. This clearly involves the complex notion of consciousness, (what do we mean when we say that learners 'see' something in the language?) of which more below in section 3.4. For the purpose of explaining the nature of the 'weak' interface between implicit and explicit knowledge, we need only note that for implicit knowledge to be gained through explicit instruction, it is claimed that the learner needs to be 'developmentally ready' to acquire it according to the supposedly 'natural order of acquisition' (Peineman 1984, 1989). Implicit knowledge of structures that are not 'due' to be acquired next cannot be gained through explicit instruction, 'input enhancing' or 'consciousness- raising' (Sharwood-Smith 1981). In the 'weak interface' position, the role of explicit knowledge is to enhance the speed of a learner's progress in gaining implicit knowledge, not to direct it.

3.3 Research findings for implicit and explicit learning.

It was observed above that one of the severest criticisms of Krashen's insistence on the 'no-interface' between implicit and explicit knowledge of language structures is that it is empirically unfalsifiable. Although we can show that a learner possesses particular items of implicit knowledge about a language there is no way of showing whether this knowledge was gained through an exclusively implicit or explicit route. There is, however, a body of empirical research in cognitive psychology that has addressed the processes of implicit and explicit learning and which can help to shed some light on the process of second language learning.

There is evidence for the dissociation of implicit and explict learning from studies of patients suffering from Korsakoff's syndrome (a form of amnesia in which information in short-term memory is lost totally rather being transferred to long-term memory: these patients are unable to recall any new experience for longer than a few minutes). A classic

example was described Claparede (1911, cited in Ellis 1994) who shook hands with a Korsakoff patient while holding a concealed pin, causing her to flinch and withdraw her hand. A few minutes later (long enough for her to have lost all memory of the event) he offered her his hand again. She refused to shake it, even though she could not give any reason for her reluctance. In this case, implicit learning remained intact despite the absence of any explicit memory of the event that caused the learning to take place. Later studies of Korsakoff patients (reviewed by Ellis 1994) have shown that a severe deficit in explicit memory, as evidenced by a complete inability to remember sequences of digits, or to recognise faces, does not mean that implicit learning is at all impaired. Korsakoff patients show normal behavioural conditioning, and a normal ability to improve on motor skills (such as mirror writing) by repeated practice. But the inability of these patients to remember sequences of letters, even after repeated practice, may indicate that, for this kind of learning at least, some explicit recall is necessary.

Reber and colleagues (e.g. Reber 1976, 1989, Reber, Kassin, Lewis and Cantor 1980) have studied the differences between explicit and implicit learning of 'artificial' or 'finite-state' grammars i.e. strings of letters which are generated by certain branching or recursive patterns ('rules'). In the first of these studies (Reber 1976) subjects were trained to learn these patterns under either implicit or explicit conditions. In the implicit condition, subjects were instructed to simply memorise the strings. In the explicit condition, subjects were encouraged to look for rules. Subjects in both conditions were then tested on the 'grammaticality' of new strings of letters. The results showed that subjects in the implicit condition were better at judging the grammaticality of the new strings than those in the explicit condition, even though it was these subjects who had been consciously looking for the 'rules'. Reber concludes that implicit learning of abstract rules is thus superior to explicit learning. Interestingly, when the subjects in the implicit condition were asked to say why they judged some strings as grammatical and others as not, they were unable to give any explanation, indicating that they had gained some knowledge of the rules without ever consciously thinking about them. Reber's view is that subjects learned the rules of the artificial grammar without consciously forming and testing hypotheses about them

(because they had been instructed only to commit the letter strings to memory), but that nevertheless they had unconsciously induced them.

When this experiment was operationalised differently (Reber et al. 1980), the subjects in the explicit condition were given some training with the rules of the letter strings through diagrams and examples. Armed with this advantage they did considerably better at grammaticality judgements than subjects in the implicit condition who were only instructed to memorise letter strings. From this, Reber et al. conclude that although implicit learning is superior to explicit learning-through-rule-searching, a distinct advantage for explicit learning is gained when it is aided by explicit teaching of the rules. When explicit learning is 'primed' in this way, it produces a performance that is superior to implicit learning.

A slightly different approach was adopted by Nissen and Bullemer (1987). In their experiments subjects had to watch a row of four lights with buttons underneath. When a light came on, they had to press the button under it. For the control group lights came on in a random order and results showed no measurable increase over time for the speed or accuracy of the subjects' responses. For the subjects in the experimental group however, there was a repeating ten-light sequence. Although the subjects did not know there was a sequence, and although it was quite unnecessary to know the sequence in order to carry out the task, there was clear evidence that the speed and accuracy of button-pressing increased over time, indicating that some learning of the pattern had indeed taken place. The nature of this learning was investigated in further tests. In the first, the subjects had to answer a questionnnaire which asked them if they had noticed a sequence and whether they could describe it. This was to assess their awareness of the structure they appeared to have learned. In the second test, subjects were asked to watch the lights again and to predict which light would be next to come on. This assessed their ability to use explicitly their knowledge of the sequence. The results of both tests were highly variable, indicating that neither conscious awareness of the structure nor the ability to use knowledge explicitly is necessarily related to learning. Interestingly, Nissen and Bullemer included some Korsakoff patients in their experiment. These subjects, unsurprisingly, did not report

noticing any sequence to the lights, nor could they predict which light would come on next, but their button-pushing became quicker and more accurate over time, just as the normal subjects' did.

Such results have suggested to Nissen and Bullemer, and to others (e.g. Curran and Keele 1993) that the mechanisms of structural learning can operate independently of conscious awareness or conscious will. How much this illuminates the mechanisms of language learning, however, is far from clear. What do sequences of lights or letters have in common with language syntax and morphology? Can we generalise from one to the other? Again we are confronted with the recurrent question of how far natural languages differ qualitatively from other systems of knowledge. VanPatten (1994) and Schmidt (1994) have both cautioned against generalising from invented sequences to natural languages because there are too many fundamental differences. VanPatten lists three that distinguish artificial ('finite-state') grammars from natural ones: finite-state grammars do not have rules of movement or possibilities of embedding, finite-state grammars do not have morphological inflections or phonological adjustments, finite-state grammars do not have any referential or pragmatic meaning whatsoever. Schmidt (1994b) adds that research experiments into the learning of finite-state grammars are conducted through intensive short training sessions and learning is said to have taken place even if subjects perform at above chance levels only. This is quite unlike the learning situation of any natural language, either as an L1 or an L2, in which nativelike performance is the requirement for success, and learning occurs gradually over several years. In short, finite state grammars are a very long way from natural languages, and it is hard to see how research findings of investigations into the short-term learning of a reduced and regular 'syntax' can illuminate how implicitly or explicitly a complex, varied and communicative system might be learned over time

Other studies into implicit and explicit learning have avoided this criticism by using natural languages or theoretically grounded artificial languages. Ellis (1993) investigated the learning under three conditions of a complex morphological rule in Welsh by native

speakers of English. The 'grammar' group was taught the rule explicitly. The 'random' group was shown a series of examples of the rule in action without any explicit explanation of how it operates. The 'structured' group was given both an explanation and examples of the rule in action. All three groups were then trained in learning the English translation of the same Welsh phrases containing instances of the morphological rule. In order to test their learning of the rule, all subjects had to make judgements on the well-formedness of a series of Welsh sentences and to undergo a post-experimental 'debriefing' session that tested their explicit knowledge. The random group performed worse, showing little evidence that they had any implicit or explicit knowledge of the rule. The grammar group showed explicit knowledge of the rule, but little evidence that they were able to apply it. The subjects in the structured group performed best, scoring highest on the well-formedness test and demonstrating explicit knowledge of the rule. Ellis argues that it was the 'synergy' of explicit rule knowledge and knowledge gained from exposure to helpful examples that enabled the structured group to outperform the other two.

DeKeyser (1995) used Implexan, an artificial language,[9] to investigate implicit and explicit learning. Implicit-inductive learning was implemented by pairing pictures with Implexan sentences that described them, but without any explanation of the rules that governed the word form or word order of the sentences. Explicit-deductive learning was implemented by giving an explanation of a rule, followed by an Implexan sentence containing the rule in action, paired with a picture. As is the case in natural languages, rules were of two types: straightforward, 'categorical' rules that always applied, and less predictable 'prototypical' rules which sometimes applied and sometimes did not. Subjects were tested on their ability to produce correct sentences for a series of pictures, some of which they had seen before in the training sessions, and some of which were novel. There was very little difference (90% and 89% correct) in the performance of the two groups on providing Implexan sentences for the familiar pictures, but for the novel pictures it was found that the explicit-deductive group scored far better for categorical rules than did the implicit-inductive

[9] Distinct from an artificial grammar of meaningless letter strings in that it is composed of a meaningful vocabulary of nouns and verbs, each subject to morphological rules

group (57% correct compared to 33% correct). DeKeyser suggests that the success achieved by the implicit-inductive group on producing correct sentences for familiar pictures was due to memorisation rather than any implicit learning of rules, and this is why the group failed to generalise from the familiar to the new. DeKeyser concludes that explicit teaching of simple categorical rules is more effective than mere exposure to them. However, this study did not show that the more complex prototypical rules could also be better learned through explicit instruction. In fact, for these rules, the accuracy of the explicit-deductive group was slightly worse than that of the implicit-inductive group..

The studies above have revealed a complex picture of the nature of learning. They show that implicit (unconscious) learning resulting in implicit (inaccessible) knowledge is certainly possible in the case of structural patterns in letter strings and light sequences (Reber 1976, Nissen and Bullemer 1987), and that this implicit learning is more successful than explicit searching-for-rules (Reber et al.1980). But they also show that for complex structural systems such as languages (artificial and natural) implicit learning through unfocussed exposure is not nearly as successful as explicit teaching and/or focussed exposure to helpful examples (Ellis 1993, DeKeyser 1995). However, DeKeyser's results indicate that it is not wise to make claims about the learning of grammatical rules as if they were all equal. Different aspects of language such as lexis, syntax and morphology could be learned through different degrees of implicit and explicit processing (Schwarz 1993) though even these distinctions may not be fine enough. According to VanPatten (1994) it may be more accurate to talk about distinctions *within* syntax and morphology between, say, learning 'easy' rules (best done more explicitly) and learning 'hard' rules (best done more implicitly).

The studies reviewed above do not provide us with any evidence for the absolute separation of implicit and explicit learning, only for the possibility of dissociation of structural learning from conscious awareness. On the question of the nature of the relationship between implicit and explicit knowledge, they tend to support the position of a weak interface in that explicit knowledge can be shown to facilitate the development of

faster and more accurate (automatic/proceduralised) skills. The thorny question of the learner's conscious awareness of what he is learning is raised in many of these studies. Although we can be sure that the Korsakoff patients who were studied (in Nissen and Bullemer 1987) had no explicit recall of the sequence they had learned, we do not know if they were conscious of observing it at the time of learning. From the evidence of the questionnaires, some of the normal subjects were clearly aware of learning a pattern, others were not. But all must have noticed a pattern at some level of consciousness because all had increased reaction times showing they had learned a pattern and were anticipating the position of the next light. It is perhaps uncontroversial to say that nothing can be learned if it is not noticed at some level of consciousness, but it is highly controversial to say what that level needs to be. It is to this we now turn.

3.4 Consciousness in second language acquisition.

The central question on this issue is whether or not learners must consciously attend to the form of the language they encounter in order to be able to process it and increase their knowledge of target forms. Advocates of the no-interface position (such as Krashen 1982) maintain that conscious attention needs only to be directed to language meaning, and that the acquisition of language forms follows *sub*consciously. While Krashen allows that learners do consciously attend to language forms, either because they have been so directed by a teacher, or because they themselves have become aware of something about the form of the target language, he maintains that this consciously learned knowledge is never available for fluent target language use. Only subconsciously acquired knowledge can provide this.

Many of Krashen's critics have pointed out that in the absence of any proper definition of what is meant by 'conscious' and 'subconscious' such a distinction is invalid. Krashen's one attempt to provide an empirically operationalisable definition (Krashen et al. 1987) based on unverbalisable 'feel' and verbalisable 'rule' has been dismissed as unsound (McLaughlin 1987). But Krashen is not alone in theorising about the role of the conscious,

the unconscious and the subconscious in language learning without a robust definition of terms. This has been a widespread failing in much second language acquisition writing, compounded by the unhappy fact that in everyday language the word 'conscious' is highly ambiguous and has several senses. Schmidt (1990, 1993a, 1994) specifically addresses this problem, and concludes that it is important to avoid the potential confusion of using broad 'umbrella'. terms like conscious and unconscious. Instead, he distinguishes four main senses of the concept of consciousness. These are: *consciousness as awareness*, (which involves the implicit/explicit contrast) *consciousness as intentionality*, (which involves the contrast between intentional and incidental learning), *consciousness as control* (which involves the contrast between automatic and controlled processing), and finally *consciousness as attention* (which involves the contrast between focal attention or 'noticing' and peripheral attention).

- *Consciousness as awareness*

For Schmidt, it is clear that explicit learning of structured material is more effective than the implicit learning, indicating that there are advantages for learners in being aware of what they are learning. In the first place, learners who are aware of what they learning are better able to understand the corrections they receive from the teacher. Additionally, learners who are aware of what they learning are more likely to be able to understand the differences between their current interlanguage knowledge and the material presented to them. Consequently, they are more likely to restructure and transform this interlanguage knowledge to incorporate what has been newly learned.

- *Consciousness as intentionality*

According to Schmidt (1990) there is no doubt that incidental learning, ie. learning as an unsought by-product of doing something else, does occur and is commonplace. In second language learning, this would refer to the internalising of language forms (spelling, meaning, morphology etc) while intent upon the message communicated through them, (Krashen 1989, Hulstijn 1992). It is not, however, clear whether knowledge gained through incidental learning is of a different type from that gained through intentional learning, as argued by Krashen (1982) who views incidentally learned knowledge of the

L2 as more valuable than intentionally learned knowledge. For Schmidt it is important to bear in mind that there is no evidence to prove that learners are always ignorant of the process of incidental learning, nor that the knowledge they come by in this way remains inexpressible. Thus maintaining the distinction is not always helpful or necessary.

- *Consciousness as control*

Spontaneous, fluent language performance can be described as 'unconscious' in the sense that it does not involve the conscious and effortful retrieval of explicit knowledge. The development of 'unconscious' fluency in second language acquisition can be accounted for in three ways. In the first place there may be a simple speeding up of procedures such that declarative knowledge becomes proceduralised (e.g.Anderson 1989) or control processes become automatised (Shiffrin and Schneider 1977). For example, the halting provision of verbal inflections progresses to being more rapid and less conscious. Alternatively, an increase in fluency may be due to a restructuring of L2 knowledge (Cheng 1985) which makes it more efficient. An example of this might be a learner's perception that it is the voicing or unvoicing of the final consonant of a verb which determines whether the past tense '-ed' is pronounced /t/ or /d/, an insight which should speed up its execution. Finally, fluency can be increased by the learner using 'exemplars' or chunks of the target language, rather than processing utterances 'bottom up' from rules. These may have been perceived in the input, or generated previously by the learner through the effortful application of rules. From whatever source, they are stored in memory as intact units, and are available for use with much greater speed and much less processing than novel utterances. This is an aspect of fluency to which we return in chapter eight.

- *Consciousness as attention*

In this context, the word, 'noticing' is used to describe the allocation of attention to particular aspects of the target language. Schmidt maintains, that for learning to take place, noticing is an absolute requirement and that subliminal learning (i.e. without noticing) is impossible. Language learners may notice something in the L2 input because of its frequency, or saliency. If attentional resources are limited, then it is these prominent forms which stand the best chance of calling attention to themselves. Instruction can help learners to notice features of the L2 input (Schmidt and Frota 1986), channelling attention

to what may otherwise be undifferentiated and unnoticeable, thus making it available for analysis. Schmidt believes that what learners notice is also influenced by what they are ready to notice, i.e. what is 'due' to be acquired next in the order of acquisition (Pienemann 1984). Finally, whether features of the language are noticed by the learner depends to some degree on task demands. A learner engaged in a task which requires a lot of logical reasoning, for example, is less likely to have spare capacity to notice new features of the target language.

Such distinctions are illuminating and demonstrate that the role of consciousness in language learning is not one simple issue but several which are complexly related, and which it is useful to disentangle. If one accepts a multifaceted view of consciousness, then a complete separation of explicit from implicit L2 knowledge, of inductive from deductive L2 teaching, or of incidental from intentional L2 learning is impossible to justify. For L1 learning, however, where implicit learning is surely the only route (small children are not taught grammar and lack the cognitive skills necessary to induce grammatical rules) the best explanation is that implicit and inaccessible knowledge is gained not through inductive hypothesising but through memory and the gradual accumulation of associations between frequently co-occurring features[10] .

Schmidt's discussion of the role of *consciousness as attention* in second language learning has important implications for pedagogy. If noticing is a 'necessary and sufficient condition for the conversion of input to intake' (1993a) then it is presumably advisable for the classroom to be organised in such a way as to enable 'noticing' to take place. But the conundrum for L2 learning, (unlike the learning of other cognitive skills) is the competition for attention between the form of the language and the meaning it conveys. VanPatten (1990) has demonstrated that language learners, especially beginners, have great difficulty in attending to the form and the meaning of input language at the same time. His results show that learners pay attention to language form only if this is necessary

for them to be able to understand the content. If they can understand the content without attending to form, they will not attend to, or 'notice', form. It is clear that a teaching methodology that attempts to focus learners' attention onto content will be doing so at the expense of attention to form, and thus increasing the likelihood that language forms will be ignored. Equally, a methodology that concentrates upon language form will be taking attention away from meaning with the risk of isolating the knowledge of form from the ability to communicate meaningfully. Manipulating attention in the classroom between form and meaning in a productive and balanced way is the subject of chapter five. Meanwhile chapter four will review how the more traditional forms of language teaching have had as their goal the automatic control of explicit knowledge, while more recent reactions to this approach have adopted instead a goal of the incidental acquisition of implicit knowledge.

[10] It is possible, for example, that native speakers of French have no inducted rules about assigning gender to nouns, but have instead a vast memory for which article or adjective occurs with which noun. (Schmidt 1990.148)

Chapter Four

The Practice of Second Language Pedagogy.

Every style of language teaching is informed by the expectation that the methodology is beneficial to the learning process. This expectation may never be articulated, but it is hard to imagine a dedicated teacher who didn't know or didn't care to know if her classroom practice was in fact helping the students to learn. Apart from the cynical, who should probably be in other jobs anyway, language teachers assume consciously or otherwise that what they are doing is providing the environment, materials and information necessary to someone learning another language

In this chapter, we shall look in some detail at how beliefs about the way languages are learned have an influence on the way they are taught. We shall also be considering how difficult it can be to dislodge some long-held beliefs from the classroom even when research evidence strongly suggests that these beliefs are misplaced. Research into second language acquisition is a relative newcomer to the scientific field, having been active for only the past thirty years of so[11], and in some respects has been slow to percolate down to the classrooms where second languages are taught.

4.1 The Influence of Behaviourism

Until the late 1960s there had been very little empirical inquiry into second language acquisition, and there was an assumption that it could be accommodated by general (i.e. behaviourist) learning theory (Skinner 1957). It was believed that language acquisition,

---

[11] As opposed to theories of second language acquisition, which have been around for a very much longer time (Thomas 1998)

first and second) was simply a process of habit-formation, and that the errors so universally observed in second language performance were the result of old habits from the L1 interfering with the newer, more fragile processes of the L2. In other words, the L1 was a hindrance to the development of the L2, getting in the way and making progress difficult.

One age-old method to build up new and appropriate habits was rote-learning. Generations of children (including me) learned Latin noun declensions and verb conjugations in the very same way that they learned their times tables.

amo	*I love*	amavi	*I have loved*
amas	*you love*	amavisti	*you have loved*
amat	*he loves*	amavit	*he has loved*
amamus	*we love*	amavimus	*we have loved*
amatis	*you love*	amavistis	*you have loved*
amant	*they love*	amaverunt	*they have loved*

These could be set for homework and successful learning could be easily displayed by recitation or drills.[12] Errors meant that the learning had been insufficient, and the paradigm had to be memorised again.

The learning of Latin, at least in the twentieth century, did not usually contain any element of speaking, confining itself to the written word (and very often only to the translation of the written Latin to the written mother tongue). Modern Languages are, of course, taught not so much for translating histories or poems, but to enable the learner to understand and produce the spoken L2. For this, other teaching techniques are required. The audio-lingual method of language learning, which was widely promoted in the 1960's and is still confidently marketed in the print media today, concerns itself very much with the spoken

[12] Learning certainly is very successful this way: I have no more forgotten my declensions and conjugations than I have forgotten my times tables, even though I have used Latin far less often.

word. It is based (loosely) upon the behaviourist idea that language skills are learned through automatising controlled processes (see for example, Rivers 1964). Learners listen to an audiotape of a series of target language prompts and provide the correct responses. Immediately after the learner has given a response, he hears the correct form on the tape and is able to judge whether his own response was accurate or not. If it is not accurate, or not fast enough, then his learning is insufficient and he has to listen through the tape exercises again. Alternatively, the learner is drilled in providing the correct responses to stimulus prompts by a teacher who immediately praises or rejects the learner's utterances and who will repeat the drills until the responses pick up speed and accuracy. In either way, this basic stimulus-response-feedback methodology is supposed to enable the learner to develop good language 'habits' that will result in his becoming faster at supplying the correct responses, and ultimately to his being able to reply effortlessly and instantly. Inaccurate or slow responses mean more practice is needed.

Such a quintessentially behaviourist approach is very close to the Shiffrin and Schneider view of learning through automaticisation which we examined in chapter two. There an automatic process was described as a 'learned sequence of nodes' from the long-term memory store that is always activated and moved into the short-term memory store in response to a particular stimulus, without the need for conscious attention or control by the subject. It is obvious, however, that such a view of learning is not adequate for language because it is more than automatically supplying a set response to a familiar stimulus. It involves being able to understand utterances that are unfamiliar, and being able to form appropriate responses that conform to the syntactic, semantic and phonological rules of the L2. Advocates of the audiolingual (e.g. Brooks 1960) tend to be a little vague on this matter, but assume that learners will implicitly acquire knowledge of structures through sufficient exposure to target language stimuli, and this knowledge will enable them to become productive users of the language.

4.2 The explicit teaching of grammar.

The debate about implicit and explicit learning was discussed in chapter three above and need not be reviewed here. From the teaching perspective that we are taking in this chapter, we need only observe that teaching a language through explicitly teaching its rules is a methodology with an ancient pedigree that still flourishes all over the world. For many learners and teachers, a language is, quite simply, *the sum of all its rules.* Lessons might consist of teaching a rule, and using it to help break the code of the written L2. Or, more recently with the advent of communicative language teaching, a lesson plan might be built around a particular rule which will first be carefully explained, then practised in role-plays or discussions so that learners can get quicker and more confident in processing it. (See the PPP approach discussed below in 4.6.1). Syllabuses can be straightforward compilations of structures, with 'easy' ones taught first, and more complex ones taught later. Examinations can be straightforward also, requiring candidates to demonstrate their knowledge of the L2 grammar (probably expressible only in the L1) through transformation or translation exercises.

In this view explicit knowledge of grammatical structures is the bedrock of performance, and must be sound. When errors occur, they are seen as the result of insufficient understanding or poor learning, requiring the rule to be explained again and perhaps more practice given in using it. A good teacher is one who knows all the rules and can explain them simply and convincingly. A good learner is one who can pick up rules and apply them accurately in fluent speech. In contrast to the rote learning of verb and noun morphology discussed above (which was seen to have much in common with the Shiffrin and Schneider model of learning through automaticisation), the learning and application of rules is more like Anderson's ACT* model (1982, 1983, 1992) in that conscious declarative knowledge is fine-tuned and proceduralised, and its execution becomes less demanding of attention.

Whether taught by rote or by rule or by a combination of both, successful language learning is most often measured by the absence of errors. From the behaviourist point of

view, the absence of errors indicates that the learner has overcome the interference of L1 habits and formed correct L2 habits: in other words the absence of errors indicates learning.

4.3 Contrastive Analysis

Being able to predict where the L1 is likely to interfere with the L2 could be a very useful tool for teachers because it would enable them to concentrate their teaching most efficiently. For some years from the late 1950s to the early 1970s a body of research which came to be known as Contrastive Analysis (see for example Lado 1957 and Lee 1968) claimed that observed differences in the grammar, lexis and phonology between the learner's L1 and L2 could be used to predict where L2 errors would occur. These areas could then be targetted by teachers in order to give intensive practice in forming new and correct L2 routines, overcoming the otherwise inevitable transfer of inappropriate L1 routines. Equally, the places where an L1 and L2 were very similar in grammar, lexis and phonology could be highlighted as areas where L1 transfer is actually very beneficial, providing unproblematic and speedy progress in the L2.

Although a considerable amount of work went into describing and comparing different languages in order ascertain where learners would make errors (e.g. Stockwell and Bowen 1965), Contrastive Analysis came under increasing criticism through the 1970s. Doubts about the validity of the approach arose when researchers began, for the first time, to look closely at the language actually produced by L2 learners. Empirical studies by, among others, Grauberg (1971) George (1972) and Dulay and Burt (1973, 1974) revealed that the great majority of morphological errors in learners' performance could not be simply explained as arising from L1 interference. In these studies the very practical problem of coding errors according to their psycholinguistic origin was clearly demonstrated. How can one reliably tell if an L2 error is reflecting L1 structure, or if it is reflecting a stage in language development as observed in first language acquisition? Many errors are ambiguous in origin, others cannot be sourced at all. (See also Felix 1980). But even

allowing for the difficulty in assigning errors to categories, the empirical studies of the early to mid-70s all showed that L1 interference is most certainly *not* the cause of between two-thirds and one half of errors in L2 performance. These must therefore be coming from some other source.

While this empirical research showed that contrastive analysis is an unreliable predictor of most L2 errors, let alone all of them, other theoretical criticisms of the approach were gaining strength. In particular there was an objection to the equation that Contrastive Analysis seemed to draw between *difference* and *difficulty*. The existence of a difference in, for example, the way an L1 forms interrogatives and the way an L2 forms interrogatives does not necessarily mean that the learner will find L2 interrogatives difficult.[13] Nor can the degree of difference in two languages be necessarily equated with the degree of difficulty a native speaker of one will encounter in learning the other (Lee 1968). In fact it can be argued that subtle differences are actually harder to learn because they are easier to miss, a point first made as long ago as 1927 by Skaggs and Robinson (James 1980). It is quite possible that someone whose L1 is Portuguese may continue to make repeated and unnoticed errors in Brazilian Portuguese which a learner whose L1 is Chinese may never make at all because the correct forms are, for him, so much more distinctive. Studies by Wode (1976) and Jackson (1981) confirmed that there was no positive correlation between linguistic difference and learning difficulty.

Another criticism levelled against Contrastive Analysis was that even if *difference* can be shown to mean *difficulty*, it does not necessarily follow that *difficulty* entails *errors.* It is quite possible for a learner to make errors in speaking without experiencing any difficulty. Equally, a learner can produce a beautiful target language sentence free of any error, but this cannot be taken to mean that he did not experience difficulty in producing it. There is no reason to suppose that there is a causal relationship between difficulty and error.

Thus Contrastive Analysis came under fire. Though it held out the promise to teachers of being able to understand where their learners' errors were coming from, in practice it could do this only to a limited extent. Nor could it help much in predicting what parts of a language a learner was likely to find difficult, or where errors were likely to occur. Worse, Contrastive Analysis ignored factors such as attitude, aptitude and other individual differences in learners which research was beginning to show have an important influence on learning (Skehan 1986, 1989). As a guide to teaching and syllabus design, Contrastive Analysis has largely disappeared from view, though some (e.g. Shortall 1996) regret the loss of what can sometimes be a useful tool for language teachers. Clearly such a thing as L1 transfer does occur, and teachers can certainly benefit from an understanding of this source of their students' difficulties.

Error Analysis, on the other hand (Corder 1967, 1971, 1974) sought not to show that an L1 interfered with the learning of an L2, but to see what information L2 errors could give about both the current state of a learner's knowledge of the L2 and the strategies learners employ to simplify their task. An Error Analysis is done on a homogeneous sample corpus in which all the errors are identified, classified, explained and evaluated. From this information it is possible to see what kinds of linguistic errors appear in particular samples, and to see that strategic devices (e.g.overgeneralisation of morphological rules, over-extension of word-meanings and a preference for simple syntax) are commonly adopted. Although Error Analysis has not illuminated a 'natural route' followed by all L2 learners, it has shown that errors should not be seen always as the result of non-learning, or predictable L1 interference, but as 'a guide to the inner workings of the language learning process.' (Ellis 1986: 53). Accepting errors as a necessary and unavoidable part of the language learning process represents an important shift in attitude, as we shall see in the sections that follow

---

[13] There is, for example, no evidence that Japanese learners are apt to produce English interrogatives like, 'You on the beach at volleyball have played?' even though a contrastive analysis of the word order of

4.4 Developmental language learning

This shift in attitude had its origins in the field of first language acquisition. A decade of research from the mid 60's to mid 70's was accumulating evidence that strongly challenged the classic behaviourist account of L1 acquisition. As the utterances of young children were collected and analysed, it became obvious that the majority bore no resemblance to the kind of utterances that these children actually heard from their caretakers, and therefore the argument that L1 acquisition proceeds by an imitation-reinforcement mechanism was untenable. Longitudinal studies of young children learning their first language showed that its development is continuous and incremental and can be characterised as a series of stages. The first utterances of infants are holophrastic (one word). Later these are replaced by two-, then three- and four-word utterances as the child's grammatical knowledge and memory capacity increases. Grammatical knowledge itself is built up in a series of stages. Complex structures are mastered gradually over many months as the child appears to operate a system of hypotheses under constant revision. For example, in learning to use negatives in English, a child will typically begin by using the word 'no' to negate everything:

no juice

no walk

no up.

no red

Then 'not' is used to negate verbs, adverbs and adjectives, leaving 'no' for use with nouns:

doggie not walking

baby not sleeping

that not pretty

no shoes here.

Later, auxillary verb contractions are used complexly and correctly to indicate both tense and negation:

Mummy isn't sleeping now

Japanese might suggest they should.

the car won't go anymore

I didn't do it.

It is not difficult to see how this process of hypothesis-testing and refinement could be applied to second language learning. As we noted above, Corder (1967) was one of the first to suggest that errors were as much part of the second language learning process as they were part of the first language learning process. Second language learners do not master target language forms in one great step, but many small ones. Their progress mirrors the way young children gradually acquire the grammar of their mother tongue. In fact the stages illustrated above of the way children learn the complexities of English negatives can be paralleled in the way many learners of English produce negatives over a period of time:

no like coffee

I no like coffee.

I don't like coffee.

Corder (1971) referred to these stages in learning as 'idiosyncratic dialects' that were part of an evolving 'transitional competence'. The term that has come to be widely adopted for this phenomenon, however, is 'interlanguage' (Selinker 1972). This notion of a continuum of L2 learning involving successive refinements of a grammatical system through hypothesis-testing recasts L2 performance errors in a much more positive light. No longer evidence of poor learning, errors are seen as manifesting 'learning-in-progress'. As we shall see below this has very important pedagogic implications.

4.5 The 'Natural Order' of language learning

If L2 learners are progressing along an interlanguage continuum, are they all passing down the same route, and is this route the same one that all L1 learners pass along? Research into the possible 'natural order' of acquisition of English negatives (summarised by Schumann 1979), English interrogatives (e.g. Butterworth and Hatch (1978) Wode (1978)) and English relative clauses (e.g. Schumann 1980) has provided good evidence

that there is indeed considerable similarity in the way learners with different L1s develop their control of these structures, though this similarity should not be overstated (i.e. there are broad stages that everyone passes through, but some of the smaller steps may be left out or else learned in a different order and at a different rate). There is also some evidence that this sequence of development is common to both L1 and L2 learners, suggesting that the two processes are operating in the same way, i.e. there is an internal syllabus, determined by the operation of Universal Grammar which all learners naturally follow. Using further research to determine more clearly what this internal syllabus consists of for a given language might give teachers a better idea of how to organise a teaching syllabus for it, but so far the research results have not been altogether convincing (Peineman 1984). As we shall see below in section 4.7, the idea that there is a natural route for the learning of an L2 has given rise to a school of thought in language teaching methodology that the most effective way of teaching a language is not in fact to teach it, so much as to *assist in its growth*. In this view there is no more use in teachers focussing learners' attention upon a syllabus of L2 rules than there is in parents focusing their young children's attention upon the grammar of their L1. But before considering the classroom practices of this particular approach, however, we shall look at the effects that SLA research has had on more mainstream teaching.

4.6 Implications of SLA research findings for the classroom.

If research findings are indeed reflecting the nature of first and second language learning, the implications for the second language teacher are profound. The teacher's historic role of choosing what is to be taught, and when it is to be taught, is undermined if the learner is following an internal syllabus impervious to outside influence. Even more significant perhaps are the implications of interlanguage studies upon teaching: if learners do not learn the target forms in one step but through many stages, is there any point in teachers continuing to coax perfect target language forms from their students, rejecting and correcting any that are not target-like?

Thus research findings over the last twenty-five years or so clearly challenge the traditional methodologies of language teaching, and in response there have been changes in classroom practice that have attempted to take into account the better understanding of the way people learn languages. The teaching and syllabus writing profession has, however, been somewhat conservative and selective in what research they have taken on board. In this country and the United States, (but not in many other places in the world) there has been an almost universal acceptance that behaviourist approaches to language teaching are more likely to produce learners who can do drills than learners who can speak the target language extemporaneously. There is also recognition that explicit teaching of rules often results in learners who can explain many more rules than a native speaker could, but who cannot use these rules to produce fluent language. In response to this, language teaching in places like the UK and the USA has shifted the balance away from a concentration upon grammar towards an appreciation that language learning is best accomplished when it can be seen to help learners express 'real' meanings, relevant to themselves. Communicative Language Teaching (CLT) methods use 'authentic' examples of the target language and attempt to engage the learners in communicating about subject matter that is useful, amusing, interesting or controversial. Fluency and communicative effectiveness are prized above grammatical accuracy.

In its strictest form, CLT avoids any explicit teaching of isolated language structures. However, it is probably true to say that in general CLT methods have not replaced structural teaching, but have been adopted alongside it. There has been no wholesale rejection of rule-based teaching, with its drills and memorisation of conjugations. Instead there has been a skillful marrying of these old methods with more innovative communicative activities such as role plays, games and group work.This compromise of the old and the new is typified in what is known as the PPP approach, currently the most influential (i.e. most commonly used by teacher trainers) in the UK and USA.

4.6.1 The PPP approach

The PPP name stands for presentation, practice and production. In this, the teacher (or more probably the syllabus designer) selects a grammatical structure for the focus of the lesson. This structure is presented to the class, often in the form of a paradigm, and its rules and applications are explained. The next stage is characterised by carefully constructed exercises, such as drills, that give the learner the opportunity to see the structure in action, to understand its workings, and to practice using it in a controlled and low-risk environment. In the final production stage the learner is encouraged to produce the target structure more independently and more meaningfully through, for example, pair-work, role-playing or discussions.

An example of this is the following lesson plan taken from the intermediate level of the Headway series of English language textbooks (Soars and Soars 1986). In unit 5 the structure for study is the future 'will' contrasted with the future 'going to'. The students look at a reproduction of a scribbled shopping list, and listen to a few lines of dialogue on a tape in which two people discuss what one of them is going to buy. (This is an everyday situation of relevance to almost everyone). The grammar of these two forms of the future is then carefully and explicitly discussed, with the teacher asking many questions of the class to ensure that they have understood the distinction.

> Why does Peter say:
> **I'm going to** buy some sugar: but **I'll go** to the baker?
>
> What's the difference between **will** and **going to** to express a future intention?
> (p.25)

For future reference, these differences are explained simply on the facing page of the students' book:

Will: Uses

We use will:

1.To make a future prediction (This can be a personal opinion.)

I think it will rain tomorrow.

2. To talk about a future fact.

The Queen will open the new hospital on Thursday

.

3. To express a future intention or decision, often made at the time of speaking.

A: Did you know that John's in hospital?

B: No, I didn't I'll go and visit him this afternoon.

Going to: Use

We use going to to express a future intention, plan or decision thought about before the moment of speaking.

When we go to France, we're going to stay in a hotel.

(It's already booked.)

This presentation stage is followed by a practice stage in which pairs of students are told to imagine that it is Christmas and they have to buy presents for everyone. (Again, this situation is one that most if not all learners will recognise as a normal part of life.) They are given a model:

Student A. You have already decided what you are going to buy.

Student B. You are looking for suggestions.

B *What are you going to buy for Henry?*

A *A record.*

B *What shall I buy him? What does he like doing?*

A *He likes reading.*

B *Right. I'll buy him a book.*

Having seen this example and maybe listened to a few students demonstrate the dialogue, the students then work in pairs through a list of other friends;

Anne------likes gardening

John------likes painting

Aunt Sally----likes cooking

Uncle Bob----likes model railways

Kate (age 3)---likes playing with dolls.

After this low-risk, pre-structured drill, the students are told to discuss the Christmas presents that they are going to give or will give to other members of the class and members of their family. It is this last 'free production' stage which is most properly communicative. The learners are given the freedom to use the target language in a way which is particularly meaningful to them without the constraints of the textbook prompts. In other examples of the production stage they may have to talk about themselves with a partner about where they live and what they do. Or they may have to imagine themselves in a particular situation, such as asking for directions, buying tickets for a train, or apologising for being late for work, and to act out with a partner what they would say in these circumstances. In this way the language under study is put into context, by attempting to demonstrate its use (and usefulness) in situations that are likely to be encountered in real life.

The PPP approach can be seen to have taken on board the psychology of communicative language teaching that students learn better when the language they are studying is relevant, interesting and obviously useful. This is clearly a welcome change from the days of Stannard Allen's *Living English Structure* (1947). This popular book of grammatical explanations and transformation exercises went through several editions and was still in widespread use in the 1980's, even though its examples of the English language were very remote from the life experiences of most language students (and language teachers).[14] In addition to being communicative in nature, the PPP approach can also be seen as mirroring the acquisition of skill as described by cognitive psychology: conscious declarative knowledge is applied through slow and controlled practice so that faster procedural knowledge can be developed.[15] But as for recent research into the way first and second languages are learned developmentally through internal processes, not a great deal of influence can be detected here. It is apparent from the Headway lesson plan discussed above that there is still a strong assumption that languages are learned through an explicit understanding of their grammar, and this forms the backbone of each lesson unit. From the accompanying teacher's book (Soars and Soars 1986b), it is clear that errors are treated as evidence that a learner's grammatical understanding understanding is insufficient:

> (Students') production of English, both spoken and written, needs to be diagnosed for error, so that areas of weakness can be pinpointed and

---

[14] Consider for example the following sentences from Stannard Allen used for an exercise to turn the active to the passive voice:
They tell me someone has shot your uncle.
One cannot eat an orange if nobody has peeled it.
Ladies usually go to a tea party more to speak to other people than for other people to speak to them.
[15] It would be wrong, however, to conclude that positive research findings in psychology are responsible for the enduring popularity of the PPP approach. The vast majority of teachers and trainers are probably happily unaware of what cognitive psychology has been up to in recent years

> worked on. Language items that perennially cause problems need to be revised. (Introduction p 1)

In terms of book sales and number of teachers trained to use this method, the PPP approach is hugely and enduringly successful. But in accounting for this success it would be wrong to suppose this is because it is so good at turning out competent L2 speakers. In fact, in this country the levels of attainment in foreign language learning through this methodology (as well as through others) remain generally disappointing. My own language teaching experience over eighteen years, and that of my colleagues in the United States, Japan and the UK, is probably mirrored in classrooms everywhere: many language students do not significantly improve their proficiency no matter how many times they retake courses.

Failure to progress in a foreign language is usually treated not as a sign for redesigning the teaching technique, but as a sign for giving the learners more of it, i.e. declarative knowledge of certain rules has to be relearned, there needs to be more drilling in order to develop more automatic control of structures, there needs to be more opportunities to use the target structure in free production. Although there is good research evidence pointing away from teaching methods and towards individual differences (e.g. in motivation and aptitude) in the learners themselves as causes of successful SLA (Skehan 1989), teachers and textbook writers alike generally have more faith in teaching as the most important factor in success.

The staying power of PPP teaching is not unaccountable. Its great popularity is based on several considerations that have less to do with SLA and more to do with teaching convenience. These are the ease with which the methodology can be taught to trainee teachers, the simplicity of a typical lesson plan, the neat and straightforward way in which it can deliver a structural syllabus piece by piece, and the power it offers teachers in determining what should be taught and what should be tested. These are important factors and have, according to Willis and Willis (1996), given this methodology an almost

unassailable position in language textbooks and language classrooms across the world, despite the good research evidence, outlined above, that languages are not learned in the same way that Lego models are built: each piece firmly in place before the next one is fitted on.

4.7 Input and SLA

Although a lot of research into second language acquisition is carried out with no thought whatsoever for the implications it might have for pedagogy, in the seventies and eighties a number of researchers directly addressed themselves to classroom practitioners in order to advise on best teaching practice. Most notable of these was Stephen Krashen (1981, 1982, 1985) whose influence in many quarters has been considerable. In this section we shall look at Krashen's ideas and the kind of language teaching methodology that he advocates.

Central to Krashen's approach is the distinction he makes between language acquisition and language learning.[16] Language learning is the result of conscious study of the structure and vocabulary of the target language and is guided by attention to form and error correction. Language acquisition, on the other hand, is the unconscious process that we all go through as young children beginning to speak our first language. This is done not by conscious attention to rules but through engaging in meaningful interactions with other speakers. Krashen is insistent that fluent and natural performance in the L2 is only possible through acquired language, and that 'learned' language cannot become 'acquired' language, (a contentious issue in SLA which we discussed in chapter three above), with the obvious conclusion that the explicit teaching of L2 structures is not particularly useful.

Krashen uses evidence for the natural order hypothesis to support his rejection of the explicit teaching of language structures. As we saw in section 4.5 above there is a body of research findings indicating that language rules are acquired in a predictable sequence that

---

[16] For many in the field these terms are interchangeable, and the attempt to draw a distinction between them is fraught with difficulty.

has nothing to do with the order in which they might be taught. It is therefore no good imposing an external syllabus on the learners: all you will achieve this way is learned knowledge of rules that cannot be put to fluent or intuitive use. For fluent and intuitive use, languages must be acquired along the natural order of the internal syllabus. According to Krashen, this is achieved by learners receiving plenty of *comprehensible input*, or in other words, by learners understanding what is said or written in the L2. Successful teaching, therefore, must be built around contriving samples of the L2 that the learners can understand. For learning to take place, the input that the learners are exposed to must contain forms and structures in the L2 that are just beyond the learner's current proficiency (the so-called 'i+1' level) and which, according to the natural order of learning, are therefore 'due' to be learned next.

In addition to receiving sufficient amounts of comprehensible input the learner needs to be in a situation where he feels confident and free from anxiety. Anger, boredom and self-consciousness are emotions that can impede language acquisition by raising what Krashen calls an Affective Filter. If the learner's state of mind is affected by negative emotions such as these, his mind screens out the target language input, even though it might be at the correct level of comprehensibility, and makes it unavailable for acquisition.

The message for teachers then is that they should provide lots of properly tuned input, and a relaxed unthreatening classroom. In such a situation language acquisition should be able to proceed down its natural route, untroubled by the learning of explicit rules and constant error correction. A relaxed unthreatening classroom is exactly what Communicative Language Teaching is good at creating. Learners are encouraged to talk to each other in group- and pairwork, the teaching materials are designed to be as relevant, interesting and engaging as possible, and there is room for lighthearted activities such as games, quizzes and music.

Krashen had been criticised for his insistence that learning cannot turn into acquisition and therefore should not be encouraged. Nevertheless, although it has been pointed out many

times (see McLaughlin 1987 for example) that it can never be proved that fluent L2 performance is the result *only* of implicit and unconscious acquisition, there is still a fairly robust school of thought among some foreign language teaching theorists and practitioners that wherever possible language rules should not be taught explicitly but be allowed to take shape in the learner's mind.[17] Krashen has also been severely criticised for his recommendation that learners have to be provided with input that is tuned to the correct level of comprehensibilty. This is a difficult matter. How can a teacher know if the samples of the target language she is providing to the class is really at the required 'i+1' level, and not 'i+3'.or 'i+4'? And how can input be tuned to fit everyone when, as any teacher knows, classes are not made up of learners all at precisely the same stage in L2 acquisition? Moreover, Swain (1985) has shown that students in the French L2 immersion programme in Canada continued to make errors in French language despite receiving seven years of schooling. The input they had received was clearly comprehensible because they achieved similar scores in mathematics, history, geography etc. as anglophone students taught through the medium of English. Swain argues that comprehensible input is demonstrably an insufficient condition for progress in SLA.

4.8 Interaction and SLA

Proponents of the view that interaction is a vital part of SLA (for example Long 1985) accept Krashen's claim that comprehensible input is central to language acquisition, but do not believe that it is difficult to fine-tune input to the right level for particular learners. They maintain that input which is too hard for learners to understand (that is, input which is at a level higher than i+1) is not necessarily lost. Instead, it can be made comprehensible through interactional adjustments either between learners themselves or between learners and native speaker interlocutors. Long and others argue that native speakers consistently modify their language when talking to non-native speakers in order to make sure that they have been understood. They may speak more slowly than usual, choose less complex

---

[17] See for example the language teaching textbook produced by Brumfit and Windeatt (1985)in which abstract diagrams of grammatical constructions are used instead of explicit verbal explanations.

structures than is normal, and repeat or paraphrase utterances. They may also use comprehension checks (such as, *'do you understand?'*) in order to determine if the non-native speaker has followed what they have been saying. For their part, non-native speakers attempt to keep track of what is being said to them by making clarification requests (such as, *'what do you mean?'*) and confirmation checks (such as, *'do you mean x'?).* In this way the input reaching the learners is modified, phonologically, semantically or morphosyntactically, in ways that keep it at the optimum 'i+1' level. The argument is thus that as interactional adjustments make input comprehensible, and as comprehensible input promotes second language acquisition, then interactional adjustments promote second language acquisition.

For advocates of Communicative Language Teaching, the idea that interaction promotes SLA is welcome news as it gives more reason to encourage learners to talk to each other or to the teacher through the use of communicative activities such as games, role-plays and discussions. In the mid-eighties there was a considerable amount of research (e.g Varonis and Gass 1985, Gass and Varonis 1985, Doughty and Pica 1986, Rulon and McCreary 1986, Pica et al. 1989; see also Pica 1994 for an overview) which suggested that information-gap tasks give the most opportunities for the 'negotiation of meaning' i.e. checking and clarifying of interlocutor utterances. Information-gap tasks are the kind in which information needed by everyone is held by only one person who then has to transfer this information fully and clearly to the other participant(s). A typical information gap tasks is 'Spot the difference' where two learners are given different copies of a picture and have to find out what exactly the differences are without showing each other their version. In a task such as this information passes both ways between the participants. In other information-gap tasks the information may only have to go one way, as in certain drawing tasks where one participant has to describe a diagram to a partner so that the partner can reproduce it exactly.

There are many versions of information-gap tasks, but each has the same basic rationale: hide certain information from one or more participants so that, in order to get it, they need

to understand precisely what their partner is saying. This will mean input is kept comprehensible and, if it contains forms and structures which are just a little beyond the learner's current level of competence (i.e. the crucial '+1') then SLA is advanced. Moreover, according to Swain (1985) SLA is also benefitted by the learner's having to modify utterances for which an interlocutor has requested clarification. This prompts the learner to pay attention to language forms and to be precise about phonology, lexis and morphosyntax.

All in all, the negotiation of meaning brought about by learners' interacting in classroom tasks is seen as an especially effective way of driving forward the learner's internal SLA processes. In contrast to a structural teaching syllabus, in which the teacher chooses a grammatical feature as the focus of a lesson in the hope and expectation that all the students will be learning what she is teaching, the negotiation of meaning promises a form of learning tailor-made to each student. There is no attempt to 'seed' an information-gap task with structures deemed ripe for attention and learning. The teacher has no part at all in determining where a learner's attention should be directed. This depends entirely on where communication breakdowns occur during the task and how they are resolved. As Skehan (1993.1) puts it, they can be described as 'the pressure points for language change.'

4.8.1 Some reservations on the role of the negotiation of meaning.

Although research has shown that a learner's comprehension is indeed aided by interactional adjustments (for example, Pica, Young and Doughty 1987) there is as yet no direct evidence linking progress in SLA with the negotiation of meaning, and it is not at all clear that it will ever be possible to separate this one form of interaction from all others in order to study its effects. In addition to this very practical problem, there are other reasons to be cautious about filling classroom time with information gap tasks. Firstly language need not be well-formed in order to provide comprehensible input. Language works very well even in poor shape because many grammatical inflections are redundant and context

can often substitue for syntax[18] . It is unlikely that missing or mistaken morphemes that should be marking plurals or verb tenses are going to cause serious communication problems and become the focus of negotiations of meaning. Secondly, people are naturally reticent in owning up to not understanding something which another person has expected to be understood. This is true for native and non-native speakers alike. No-one wants to keep saying '*pardon?*'. Related to this is a third point: people feel understandably irritated when asked constantly to repeat or explain something. As Aston (1986) has pointed out, a task which has been designed to maximise such episodes is likely to result in annoyance and demotivation all round. If Aston's criticisms are right it is not surprising that a classroom study into the negotiation of meaning (Foster 1998) showed very little of it going on, even during tasks which required information to be transferred. This suggests that gaining comprehensible input through checking and clarifying is not a popular strategy even in situations designed to promote it. Significantly, the study also showed (in common with Pica et al. 1989) that where the negotiation of meaning does occur it is more likely to result in lexical than morphosyntactic adjustments, i.e. the theoretically most valuable type of negotiated language is the least common because it is a gap in lexis, not in syntax, that is the most frequent cause of communication problems. Teaching methods which seek to maximise communication breakdowns in order to harness the benefits that arise from their resolution are probably neither popular nor effective.

4.9 Task-based learning

In the PPP approach to language teaching discussed above in section 4.6.1, classroom tasks are largely confined to the third 'production' stage of the lesson. As we have seen, this approach is fundamentally structuralist in that each lesson focusses upon a particular grammatical item and aims to deliver it to the learners. The production stage may be a task specially designed to give learners the chance to use this structure, as in the example used in section 4.6.1. Or it may be a task which aims to give the learners a chance to speak more freely with each other, without requiring the use of any particular structure. Many

---

[18] This is, after all, the basis of all pidgin languages: lots of contextual clues, little to no syntax.

lesson plans include both kinds of production tasks, with the latter especially seen as light relief from the concentration upon grammar required by the first two stages. Advocates of task-based learning, however, reject the notion that language learning is primarily driven by the explicit teaching of grammar and want to put language tasks far more into the centre stage by making tasks themselves the basis for the whole syllabus. This section will consider the proposals for a task-based approach to language teaching mainly by looking at the nature of tasks themselves. The further problems of how to grade them for difficulty, how to sequence them in a principled way, and how to test the learning they have brought about, are complex, insufficiently researched, and beyond the scope of this chapter. We shall return to these questions in chapter five.

4.9.1 Tasks and task-based syllabuses

For Long and Crookes (1991) a task is, unhelpfully, almost anything one does in life, from tying shoes to posting a letter to painting a fence. From the perspective of a second language learner, a more practical definition is required. Yet there is no simple, widely accepted definition in use, an inconvenience to which no ready solution is at hand. In order to discuss the role of tasks in the second language classroom, however, some attempt needs to be made.

It is perhaps easier to say what a task is not than it is to define what is it. Willis (1996) has an interesting set of characteristics that define what tasks are not. Tasks do not concern themselves with language display, do not embed particular grammatical structures for analysis, do not require learners to parrot out other people's meanings, and do not require learners to produce only the linguistic output imagined by the task designer. Tasks are occasions for the learner's own meanings to be expressed in the learner's own voice.

A more positive but very similar definition is one shared (largely) by Candlin (1987), Nunan (1989), Long (1989) and Skehan (1998). For them a task is essentially an activity in which attention is focussed upon language meaning rather than upon language form, in

which there is some real communicative goal to be achieved and which has an obvious relationship with the world outside the classroom.

At once we can see that many of the activities described above in behaviourist and PPP classrooms would fall foul of these guidelines. Neither drills, rote repetitions, prompted dialogues nor texts designed to show a structure in action would count as proper tasks. For Willis, what is really important is that learners have the chance to put their own meanings into words, and that they have a real (as opposed imagined) communicative purpose in doing so. There is no place in her syllabus for learners to carry out a task just for the sake of language learning. Consequently her tasks often involve learners' talking with each other about their own lives, finding out where their task partners live and work, how they were brought up and educated, what they think about certain issues. For Long and Crookes (1991) tasks should have a pedagogic goal, meaning that a needs analysis of a particular group of learners would shape the choice of tasks used in their classroom. Thus a group of air traffic controllers would be given tasks that bore a clear relationship to the work of directing air traffic. (While this may be practical for homogeneous groups of learners with a clearly defined purpose to their study, it is not obvious how practical this would be with more heterogeneous groups). Long and Crookes also set great store by the usefulness of the negotiated comprehensible input which tasks should generate, seeing this as crucial to shaping input language in ways that enable natural acquisitional process to work most effectively. Scepticism about this confidence in the negotiation of meaning, discussed in section 4.8, does not need to be repeated here.

In the Bangalore Project Prabhu (1987) attempted to develop a fully task-based curriculum in which the traditional whole-class focus on grammatical items was replaced by learners working in pairs on practical problems such as planning a journey from a train time-table, or solving a 'whodunnit'. A comparable task is first demonstrated to the whole class so that they can understand what to do and, crucially, see what language is useful or necessary to the completion of the task proper. The task proper is than carried out by the learners in pairs or small groups. Tasks are judged successful if they have engaged the

learners interest and have generated useful language. Language which is found to be useful for the completion of the task may be taught by the teacher, but there is no attempt to impose a structural syllabus on the learners. Instead they are expected to use the linguistic resources that they already possess in order to transact the task, and to use the demands of the task to develop these resources further. In other words, the assumption is that a meaningful and interesting engagement in the target language will galvanise internal language acquisition processes.

4.9.2. The problems in a task-based approach

We have already noted that it is very difficult to persuade teachers and syllabus designers to allow languages to take shape in a learner's mind without the explicit and linear teaching of a structural syllabus. Communicative Language Teaching in the United Kingdom has embraced the use of tasks but only as an adjunct to the teaching of grammar, to illustrate, to make relevant, and to enliven. The Bangalore Project, though deemed successful in its own environment, has not spawned similar projects in other countries, and the one truly task-based text book produced in the UK in recent years (Willis and Willis 1988) was not notably popular with teachers. The reasons for this are, as we have seen, partly to do with the very powerful convenience of a grammatical syllabus, but partly also they are to do with the failure of the advocates of task-based learning to convince learners and teachers alike that second languages can be learned like first languages, that is, analytically, without explicit attention to their grammatical nuts and bolts.

This is a very important quibble, born out of the observation that whereas all normal children learning their L1 proceed smoothly to the end of their learning continuum, the interlanguage of most L2 learners fails to reach target language competence. Selinker (1972) coined the term 'fossilization' to describe this cessation of learning and estimates that perhaps as many as 95% of learners fossilize varying degrees short of native-like competence. This is not necessarily a disastrous outcome of L2 learning. The world is full of people who speak second languages fluently but ungrammatically and who manage to get things done nevertheless. Schmidt (1983) reports the case of Wes, a Japanese learner

of English whose progress in the language Schmidt studied for a period of two years. In this time Wes changed from someone whose English ability was minimal to someone who was regarded by native-speakers as a highly competent conversational partner, although his linguistic competence (i.e. grammatical accuracy) did not improve much at all. Wes had become adept at using discourse skills to make inferences and predictions about the meaning of the language he heard, and strategic skills in using paraphrases and routinised chunks to express himself when his linguistic knowledge was insufficient. Wes did not take language classes because he did not care to be accurate in his English grammar, only to be able to communicate with the people around him. In this, he was very successful.

It is easy to see how classroom language tasks can create the circumstances ideal for the development of the discourse skills of inferencing and predicting without necessarily developing knowledge of morphosyntax. We have noted in section 4.8 above that language can work very well even when its morphosyntax is poorly formed. In order to understand an utterance in the target language it is not necessary for a learner to process every morpheme. Equally, in order to produce a meaningful utterance in the target language it is not necessary for him to get all the morphosyntax right.[19] There is no guarantee that learners engaged in tasks where the emphasis is very much on the successful transmission of meaning will feel the need to give attention to possibly redundant morphemes. In fact, in the light of the discussion in chapter two on the limited capacity of the brain to process parallel information, it becomes clear that the great emphasis on language meaning which task-based language teaching advocates is likely to prevent adequate attention being paid to language forms. Far from engaging internal acquisitional processes to advance SLA, task-based teaching could encourage premature fossilisation by putting learners in situations where partial understanding and poorly formed language are nevertheless successful.

---

[19] This is demonstrated in the language of telegrammes where redundant (and thus expensive) grammatical items are ruthlessly pruned without semantic consequences.

4.10 Summing up

In this chapter we have seen how second language pedagogy has, with some exceptions, resisted the influence of much recent theorising and research into second language acquisition, to remain largely guided by the belief that the explicit presentation of grammar is the safest and most accountable way to proceed. The impact of Communicative Language Teaching has been considerable in the production of teaching materials and also in the way classroom are organised, though not to the point of recreating them as places where learners come only to transact enjoyable language tasks that will engage and expand their interlanguage knowledge. Many teachers of my acquaintance remain 'closet' drillers, convinced that the only way for learners to avoid errors is to get some behavioural conditioning, even though such a belief is not acceptable to most teacher trainers nowadays. The explicit teaching of isolated language forms is still the framework of the majority of textbook syllabuses. There remains a great deal of controversy over whether errors should be eradicated by constant correction, or accepted for the sake of fluency and self-confidence as part of the learning process, with, one expects, the majority of teachers taking a commonsense rather than dogmatic view.

As we have seen in chapters two and three research into both cognitive psychology and second language acquisition is pointing away from the efficacy of explicit teaching of language rules, and there is parallel evidence from the classroom itself that teaching rules often results in learners knowing rules, but not how and when to use them in fluent language. For the learner, the tension is between prioritising language form (unnaturally slow but accurate L2 performance) and prioritising language meaning (fluent but error-filled L2 performance). There is an added tension for learners between the use of well-learned, simple forms that put no pressure on interlanguage (the 'safety first' approach) and the use of more complex, less well-learned forms that would promote interlanguage development (the 'take-a-chance' approach). The challenge for language pedagogy thus to show if a language learner can be helped to overcome his limited attentional resources in order to pay attention to both the form and meaning of the language he is attempting to

learn, with the ultimate goal of balancing fluency, accuracy and complexity in L2 attainment (Skehan 1996, discussed in 5.3 below). The task-based, 'focus on form' approach to language teaching is one attempt to bring about attention to form without compromising meaningful communication, giving planning time before tasks are undertaken is another. In the next chapter we shall look at both.

Chapter Five

Second language pedagogy: manipulating attention in the classroom.

In the previous chapter we looked at how beliefs about language learning, whether behaviourist or developmental, inform the practice of language teaching. We saw that there are two main (and opposite) approaches to language teaching: one which is built around the presentation of discrete language forms, and another which is built around a focus on language meaning. The former imposes a synthetic 'learner-external' syllabus and is delivered through such classroom practices as explicit grammar teaching, drilling, pattern practice and error correction. The latter facilitates the unfolding of an analytic 'learner-internal' syllabus and is delivered without explicit teaching of structures, through classroom tasks, especially those of an interactional nature where language meaning is the primary concern.

Advocates of one approach have tended to use the shortcomings of the other as positive evidence for their own. However, there is also interest in a different approach, the so-called 'focus on form', which seeks to avoid the shortcomings of both. It draws an important, though not always easy to maintain, distinction between the form and forms of a language. For Long and Crookes (1991: 45-46) this distinction is primarily contained in the design of the syllabus. If the content of a syllabus is the grammar of the language itself, then the classroom focus is said to be on language forms, because the learners are primarily attending to these. If, however, the syllabus is designed to teach something else, such as biology or maths, and the learners attend to language only as language problems arise incidentally in class, then the focus would be said to be on language form.

In this chapter we shall examine the practicalities of a pedagogic focus on form as a way of manipulating a learner's attention between form and meaning, and discuss research that suggests other ways of achieving this. We will also consider how a focus on form fits with the concept of information processing, discussed in chapter two above.

5.1 Focus on Form.

The idea of a focus on form as opposed to the more traditional focus on forms has been extensively written about (e.g. Long 1985, Long and Crookes 1991, Long and Robinson 1998, Doughty and Varela 1998, Doughty and Williams 1998). It is very much part of the stricter version of Communicative Language Teaching which avoids as far as possible explicit attention to isolated language structures. However, a focus on form approach acknowledges the need for classroom SLA to be helped by appropriately timed interventions, explicit instruction and negative evidence, rather than be left entirely to its own devices, a laissez-faire course which often results in incomplete language learning. (Swain 1998)

In practice, engineering a focus on form calls for particular classroom techniques. Long and Robinson (1998) suggest that during the course of a task, the teacher can guide a learner's attention to language form if problems with comprehension and/or production arise. Similarly, an information exchange task may provoke the negotiation of meaning between task participants, if they realise that their difficulties with language meaning arise from problems with language form. A focus on form may be also be brought about during 'recasts', the corrective reformulations of a learner's ill-formed utterance (often but not always by the teacher). Additionally the teacher may attempt to draw the learners' attention to something she (and not they) had noticed was amiss in their speech, though presumably without explicitly teaching the rule.

Of these techniques suggested by Long and Robinson to engineer a focus on form, only the second seems to have promise. The benefits of negotiation of meaning, as discussed

above in chapter four, are based on an assumption that communication breakdowns will occur because of syntactic difficulties, whereas there is evidence (Foster 1998) that learners might be inclined to put up with syntactic problems and press on with the interaction. Furthermore, language meaning can be successfully conveyed without the need for its form to be correct, with the result that 'redundant' morphemes (like the third person singular '-s' in English) can be omitted without any loss of comprehension. As for the teacher intervening to draw the learners' attention to a certain form which is (presumably along with other forms) being used incorrectly or inappropriately in their task interaction, this seems close to being indistinguishable from a focus on forms. It is, after all, the teacher who is choosing and isolating certain forms in order that the class will pay attention to them. Long and Robinson (1998) in advocating this as a pedagogic technique try hard to distance it from what happens in the conventional focus on forms approach, but the distinction is unconvincing.

> Having found [a particular error] to be *pervasive* and *systematic*, and (from the SLA literature and/or prior teaching experience) knowing it to be *remediable* for learners at this stage of development, he or she is usually justified in briefly interrupting the group work in order to draw attention to the problem, using pedagogical devices appropriate for students of the age, literacy level and metalinguistic sophistication concerned.
>
> (p18, their italics)

The teacher must use considerable judgement and experience to sort out which errors can 'justify' interrupting the meaning-oriented interaction of the group, and it is not clear whether Long and Robinson mean the errors of one individual or of the group as a whole. If the former, then other members of the group (or possibly the whole class) will have their attention drawn to something that is perhaps unnecessary or untimely for them as individual learners, which is a criticism levelled at a focus on forms approach. If the latter, it may be very rare for the teacher to discover an error that everyone seems to be making

and which everyone should be paying attention to. Moreover, it is not a straightforward matter, as Long and Robinson suggest, to know what errors are remediable at what stages of development.

The 'pedagogic devices' referred to by Long and Robinson which are to deliver a focus on form are not described in detail, but elsewhere in their paper (p.15) they suggest the teacher could draw attention to the learners repeatedly missing '-s' on plural English nouns by writing some on the board, underlining the 's' and saying the words slowly with emphasis on the final sound. Later (p16), they advocate highlighting some useful items in a text so that learners will notice them. Neither of these devices seems promising. The first cannot be used for linguistic items more complex than inflectional morphemes (such as the use of modal auxillaries) and the second is essentially the teacher's choice in advance of the task of what language should be noticed. In this sense, highlighting items to be noticed is akin to Sharwood-Smith's (1981, 1993) *consciousness raising* or *input enhancement* which Long and Robinson criticise for being too linguistically motivated, i.e. too much a focus on forms.

Their third suggestion, recasting learners' utterances in order to provide a timely native-speaker model, is perhaps a more promising way of introducing a genuine focus on form without interrupting a predominant focus on meaning. It allows a learner to compare a target language version with their own and thus get the opportunity to notice differences in syntax, morphology, lexis, pronunciation and/or style. Doughty and Varela (1997) report a study in which recasting was investigated as a means of drawing a learner's attention to a target language form while maintaining a focus of attention on language content. The subjects in this study were 11- to 14-year-old students in an ESL science class. The established curriculum for these students was entirely devoted to science, and no explicit teaching of English grammar was undertaken[20] While a control group of

[20] In compliance with strict Communicative Language Teaching orthodoxy, the researchers were at great pains to ensure that at no point in the study did attention to language form interrupt or distract from a

students continued to receive instruction in science and no instruction in English language whatsoever, an experimental group followed the same science curriculum with the same pedagogic techniques (hypothesising, testing and reporting results), but with the teacher consistently recasting any ill-formed past tense in both speech and writing. In this way it was reasoned that the students' attention would be drawn to the correct form of the past tense without loss of attention to the subject matter of the lesson. The results were encouraging. The students in the experimental position improved in both accuracy and total number of attempts at the past tense, especially in their oral language, whereas the control group showed very little improvement in their use of this form in either speech or writing. Moreover, the experimental group did not do any worse in the science than the control group, indicating that the linguistic focus had not been distracting the students from their main task. The researchers conclude that a focus on form, delivered in this way, is beneficial to second language acquisition, though they stress that recasting needs to be done judiciously in order not to transgress the Communicative Language Teaching prohibition on the overt teaching of linguistic forms.

It is possible, however, to see their fastidiousness in this regard as revealing the shortcomings of the approach as a language teaching method. In their study, Doughty and Varela target only the past tense forms, ignoring any other errors that the students might be making. It was the researchers who determined what the students needed to pay attention to, having diagnosed the past tense as 'task-natural' (i.e. predictably needed during science reports) and reasonably incidental to science teaching, *as well as* an area of weakness for many of their subjects. But it is not clear how, when or if other forms could be treated to a similar recasting regime within an actual content-based EFL course. Indeed this is not really Doughty and Varela's concern. Their research is aimed at showing if an implicit focus-on-form can be engineered without compromising a focus-on-content communicatively-taught EFL class, and not whether it is at all practical or sufficient for a single language form to be selectively recast by the teacher for several weeks before

---

central focus on language content (i.e. the science syllabus). It is clear that, in the view of the researchers, to have done so would have been to the detriment of the class.

moving on to recast another form, or whether several forms could be dealt with at the same time. And by strictly recasting only the past tense, Doughty and Varela may have inadvertently been reinforcing other unrecast errors in the students' speech and writing, in a sense sending the message, 'everything else is fine, just work on these past tenses'. These are important questions that Doughty and Varela leave unexplored.

Taken out of the strictly communicative straightjacket imposed by the Doughty and Varela research design, recasting is nevertheless a promising technique for guiding a learner's attention from meaning to form. When applied in a more *ad hoc* fashion, it does not require the teacher's judgement in advance as to what forms the learner needs to notice, only the teacher's on-the-spot judgement about what needs to be recast. If it is not restricted to one pre-selected linguistic item, a teacher's recast of a learner's utterance is a timely native-speaker model which might include a number of lexical, syntactic, morphological, phonological, intonational and stylistic changes, all of which become available for a learner to notice. (What actually *gets* noticed is a matter for the individual learner, and not the teacher.) It has to be acknowledged, however, that as a classroom technique, recasting is nothing new. It surely takes place in any methodology which encourages learners to interact with their teachers, just as it probably takes place whenever a native speaker is trying to be co-operative with someone learning their language, whether that person be a baby acquiring an L1 or someone older acquiring an L2.

A less positive view of recasting as an aid to language learning is taken by Lyster and Ranta (1997) and Lyster (1999). They point out that teachers do not use recasts for corrective purposes alone, but to react with approval to the content of learners' contributions, irrespective of whether these are well-formed or not, in order to continue the topic. Thus learners cannot necessarily know whether the teacher is giving feedback on the propositional content of what they have just said, or the form in which they said it. And even if learners' perceive recasts as feedback on the form of a utterance, it is not possible to discriminate between negative feedback ('*what you have just said contains an error*') or positive feedback ('*what you have just said was OK but here is another way of putting it*'). Nevertheless, in common with all utterances from the teacher, recasts provide

timely evidence for what is well-formed and appropriate in the L2. Whether teachers can intentionally use recasts as negative evidence to direct attention towards particular input-output mismatches is another matter.

In sum, a clear distinction between a focus on form and a focus on forms appears to be no easier to define in theory than it is to maintain in practice. The techniques whereby the teacher intervenes to shift a learner's attention to forms that the teacher has deemed worthy of attention would be at home in many structural, focus on forms classrooms. The idea that the teacher could decide between which forms are *pervasive*, *systematic* and *remediable* at that particular stage and which are not, and therefore which to draw attention to, and which not, imposes a burden few would wish to carry.

5.2 Planning and second language performance

In this section we shall look at a different solution to the problem of how to manipulate an L2 learner's attention between language form and meaning: giving planning time to prepare for a task. Unlike the focus on form approach discussed above, which is clearly directed at classroom practice, the notion of using planning time to affect the linguistic outcomes of a task is not explicitly offered as a classroom technique by the researchers who have explored its effects, (though clearly it could easily be incorporated into any classroom that uses tasks as part of its teaching methods). Whereas a focus on form approach requires the timely intervention of a teacher, pre-task planning is not targetted at specific moments where a focus on form is deemed useful. Instead, it is assumed that language form *in general* will receive more attention when the burden of processing language for meaning can be eased. As we saw in chapters two and three, the brain's attentional resources are limited. When taxed by processing language content (especially when this is complex or perplexing) the brain will not have sufficient capacity for attention to language form if the processing of those forms is not automatic (i.e. capacity-free), which is certainly often the case for L2 learners. Having simultaneously to process both language form and content represents a considerable cognitive burden for L2 learners. An

easing of this load might be predicted when the language content is familiar and thus less attention-demanding. Spare capacity freed up in this way could then be devoted to language form, with the result that language performance in general might be more accurate.

Research into the differences between the language used for familiar and unfamiliar topics has gone on in both L1 and L2 domains. For both, the familiarity of a topic depends on the extent to which it has been rehearsed, or planned, before being written down or spoken aloud. In chapter three we saw how a body of research in the seventies (e.g. Ochs, 1979, Kroll 1977) looked at the way native speakers coped with the demands of real-time unplanned language production on an unfamiliar topic by relying on a pragmatic mode of speech characterised in the main by simple co-ordinated syntax. By contrast, planned language was characterised by more complex subordinated constructions. We also saw that these findings were unfortunately rather speculative and based on inadequate data. However, this line of inquiry has been taken up, in a scientifically more robust way, by researchers into second language acquisition. Here the interest is in whether an effect for planning or familiarity of the topic can be shown in the language of non-native speakers, and if so, whether it is can be harnessed to increase the attention paid by L2 learners to language form, and thus to language accuracy

5.2.1 Research into planning and second language performance

Tarone (1985) was able to show that the amount of attention paid to second language speech affected the accurate use of communicatively redundant target language forms (such as the third person -s). That is to say, in a situation where communication was paramount and time to think restricted, (i.e. a conversation) accurate use of redundant forms was lower than in a situation where there was no communicative pressure and more time to think (i.e. reciting verb forms). In this analysis, attention to communicatively redundant language forms is inversely related to the amount of spare attentional capacity. Where there is insufficient capacity, there is less accuracy.

Taking his lead from the research of Ochs (1979) and Givon (1979), that planning time would increase the degree of attention paid to speech, Ellis (1987) investigated the effects of different degrees of planning on three related narrative tasks. In the first version, learners had to write down the story represented in a series of pictures. In the second version the same learners had to produce an oral version of the same story. In the third version, the learners were given a different set of pictures and had to produce an oral version of the story represented in them. Ellis reasoned that these three tasks provided learners with progressively less planning time, and that this would affect the accuracy of the language they produced, with the third version providing the least time to plan and thus the least accurate performance.

In order to measure accuracy, Ellis looked at three verb forms of the past tense: the regular past, the irregular past and the copula. His results showed that the regular past tense was most affected by the planning condition, declining from a high of 77% in the most planning time available condition (first task) to a low of 43% in the least planning time available task (third task). The irregular past tense was hardly affected at all by the planning condition, declining only slightly from 60% in the first task to 55% in the third. Ellis interprets this difference as indicating a difference in the processing of the regular and irregular past. The irregular past forms are not affected by planning time because they are accessed as lexical items and require no extra processing. The regular past on the other hand is processed by the morphological rule of adding '-ed' to the stem of the verb, and therefore takes some processing capacity. The extra planning time available in the first version of the task (writing the story down) enabled more accurate processing of the morphological rule. This result lines up with that of Tarone (1985) reported above: insufficient processing capacity means less accuracy in production.

Ellis' study has been criticised, however, for comparing speech with writing so that it is not possible to know whether the gain in accuracy was indeed due to the extra planning time available or to the fact that a written form was being processed.. Avoiding this confound, Crookes (1989) carried out a study into the oral interaction of 20 Japanese

learners of English engaged on information-gap tasks. The experimental group was given ten minutes of individual planning time before embarking on each task, while the control group had to begin the tasks immediately. Crookes compared the performance of the two groups across a variety of measures and found that the planners produced language that was significantly more complex and more varied than that of the non-planners, but not more accurate. He concluded that the learners had used the planning time to be more syntactically ambitious in the language they produced during the tasks. According to Crookes, the effect for accuracy in the Ellis study was not due to planning but to a comparison of speech with writing. His study was largely replicated by Ortega (1995) who achieved similar results.

In Crookes' and Ortega's studies it is suggested that planning time resulted in the subjects directing extra attentional resources to the complexity and variety of the language they produced to transact the tasks, though they were not able to find the extra attention needed to increase their accuracy. It is, however, possible to account differently for the lack of an effect for accuracy in the Crookes study. The tasks he used, a Lego model building task and a map directions task, do not provide many opportunities for the use of verbs beyond the simple present and imperative, with the result that it is not possible to measure accuracy in the use of, say, the past tense '-ed' or temporal auxillaries simply because they are not present in the data in sufficient numbers. Crookes makes this observation (p376) to explain why he used only the article system and plural '-s' as measures of accuracy. (Even then, he was not able to use plural '-s' as a measure of accuracy in the map task because its use here was practically zero.) His choice of the article and plural '-s' as measures of accuracy is possibly unsatisfactory because these are particularly difficult for Japanese speakers to acquire (there are no equivalent forms in Japanese) and they are both often communicatively redundant. If Crookes, or indeed Ortega who largely followed Crookes' research procedures, had used more linguistically demanding tasks, or if they had assessed accuracy through more global measures, they may have found an effect. As things are, we do not know if there was an increase in the overall accuracy of the subject's language because this was not measured.

Taken together, the research findings of Tarone, Ellis, Crookes and Ortega show the extent to which communicative pressure can affect the language performance of L2 learners. The results are consistent with the model of the brain as a limited-attention information processor which must choose between the demands of processing language for content and processing language for form. Removing the pressure of real-time communication means less attention is needed for *what* to say and more attention can be devoted to *how* to say it. The studies reviewed above suggest that this is manifested in a focus on language form, morphological or syntactic, without the need for the teacher to direct attention to particular language items.

5.3 The impact of task complexity on attention

This chapter has so far looked at two rather different approaches to the problem of how to manipulate a learner's attention. The focus on form approach requires the teacher to play a role in deciding which forms need to be focussed upon, briefly and at opportune moments, while maintaining a primary focus on language meaning. Pre-task planning is more concerned with allowing the learner to create attentional space which can then be devoted whatever aspect of language form he or she chooses. Although the pre-task planning research reported above acknowledges that the attention demanded by the content of the task is likely to divert attention away from language form, in common with the focus-on-form research it appears to assume that the content of one task will be as attention-demanding as the content of any another, that is to say, the research does not take into account how tasks may vary in the cognitive load which their contents impose upon the person who must transact them and that this in turn will determine how much attention that person is able to pay to language form.

A lack of concern for the impact of task complexity upon performance can be seen not only in some SLA research but also in the kind language testing which uses task as

vehicles for learners to display their proficiency in the L2. In all the UCLES[21] English as a Foreign Language oral exams the greatest credit is given to the candidates able to produce extemporaneous, fluent, accurate and complex language on subject matter for which he or she has had no preparation. No account is taken of how unfamiliar or complex subject matter may swallow up attentional resources. In the Cambridge First Certificate oral exam, for example, describing a completely unfamiliar photographic scene is regarded as an equal and constant task, no matter what the subject depicted, (which can be quite bizarre). Candidates who find such tasks hard are assumed to have deficiencies in the English language, rather than an inadequate imagination.

In other quarters, however, there has been concern for the way task characteristics may affect task performance. Skehan (1996) agrees that in transacting a task the L2 learner faces a tension between prioritising language meaning and prioritising language form, and that learners who prioritise the fluent encoding and decoding of meaning probably have to do so at the expense of language form, i.e. may rely on lexis and communication strategies such as gestures, facial expressions, intonation, and informed guessing. But there are also choices to be made by learners who are less concerned with fluency and who choose to prioritise language form: should they try to be as accurate as possible by avoiding complexity and using only the structures and lexis which they know well, or should they risk using more 'cutting-edge interlanguage' structures in an effort to be more complex and more target-like? Clearly, these goals are in some degree of mutual tension. It is not possible to give equal priority to each.

Though learners may be naturally predisposed to preferring to speak fluently rather than accurately, or preferring the accurate use of a small range of undemanding structures rather than adventurous use of more taxing language, Skehan believes that the tasks themselves are important factors in predisposing performance in certain directions, and that to understand how this works it is necessary to analyse and categorise task

---

[21] University of Cambridge Local Examinations Syndicate, which administers exams in English as a Foreign Language across the world

characteristics. Expanding on the ideas of Candlin (1987) he proposes two general categories of linguistic and cognitive factors and then considers how each may be further analysed into subcategories. The details of this are set out below:

1. Linguistic Factors

syntactic complexity and range

lexical complexity and range

2. Cognitive factors

Cognitive familiarity

familiarity of topic, and its predictability

familiarity of discourse genre

familiarity of task

Cognitive Processing

organisation of information

amount of computation

clarity and sufficiency of information

information type

3. Communicative stress

time limits and time pressure

speed of presentation

number of participants

length of texts used

type of response

opportunities to control interaction

(Skehan 1998:99)

In this analysis of task characteristics, the linguistic burden imposed by an individual task depends on how far it involves particular lexis or particular syntactic structures which may be within or beyond the comfortable range of the learner. Cognitively, (and perhaps more interestingly), tasks may require the learner to draw upon knowledge he or she already

has, or to come to an understanding of new information which might then need to be reanalysed or reorganised Additionally, tasks may require the learner to work entirely alone without reference to anyone else, or with one or more partners whose contributions must be taken in account. Learners unfamiliar with a particular task type may need to give considerable attention to exactly what is required of them while they transact it, especially if it involves several stages, whereas learners who are familiar with a particular task type will recognise at once what they have to do. Also important is whether all the participants in a task have the information they require to transact it, or whether crucial parts of it are missing, hidden, or unequally distributed. The level of cognitive engagement with the task material and task design can therefore be seen as attention demanding in a variety of ways, each of which could impact upon the language performance they provoke. From an information-processing perspective, it is possible that all of these factors, embracing as they do task code, content and conditions, have some bearing on how a learner's attention during a task is likely to be shared out. According to Skehan (1996), it is the job of research to discover whether there is a predictable relationship between task design and task performance which might ultimately be used in classrooms to guide learners' attention between form and meaning, and between accuracy, complexity and fluency.

5.4 Researching task difficulty

Numerous research studies over the past few years have addressed the question of what makes certain language tasks more difficult that certain others. However, before we turn to these it is as well to acknowledge a very important question:

What do you measure to find out if task A is more difficult to do than task B?

There is obviously no simple answer to this. Performance in a second language is not like performance in multiplication tables, i.e. with easily identifiable right and wrong answers. Nor is it (to use a better analogy perhaps) like reciting tongue twisters where difficulty could be measured simply by assessing speed and accuracy of articulation. Task difficulty

has to do with the amount of attention demanded of the participants. Difficult tasks require more attention than easy tasks. We know that content is prioritised over form (VanPatten 1990), so difficult content is likely to absorb lots of attentional capacity and only what is left over can be devoted to form. Researchers trying to measure task difficulty have therefore tended to reason that inadequate attention to form will be manifested in a greater degree of dysfluent (Robinson in press), or inaccurate (Skehan 1998), or uncomplex language during task performance, because these are the strategies language learners adopt when unable to give proper attention to form:

- *slow down.*
- *forget accuracy.*
- *be simple.*

While one can imagine that a task of awesome code and cognitive difficulty undertaken in extreme communicative stress might lead to the poor learner adopting all three strategies at once in order to get the task done, it is unlikely that this would be necessary for most tasks.[22] It is thus a question of which strategy is chosen by the learner to compensate for attentional resources insufficient for the task in hand. Are individual learners predisposed to an accuracy orientation, or a complexity orientation that causes them to allocate resources to one performance dimension at the expense of the other two? Or is it possible that task design could be responsible for a general shift to more accurate, or more complex, or more fluent language? Is it the case, for example, that tasks with cognitively demanding subject matter (e.g. the discussion of a particularly hot and controversial issue) will prompt learners into trying to express themselves through complex language forms because it makes them realise that simple language will not get the job done?

These are important questions which a number of research studies have begun to explore. A typical research study into task difficulty uses a design in which other variables are held constant while one task feature is manipulated. The resulting language is then measured along one (or more) of the three performance dimensions of accuracy, complexity and

---

[22] Native speakers have another strategy for situations of communicative stress: the use of pre-formed lexicalised sequences. This is discussed in chapter eight below.

fluency, and conclusions are drawn as to whether the investigated feature causes the task to be more, or less, difficult to perform. Some of these studies are reported below. They show a slowly developing picture of what the term 'cognitively taxing' might mean.

A number of researchers (Brown et al. 1984, Foster and Skehan 1996, Robinson 1996) report that tasks based on concrete and immediate information are easier (i.e. result in more fluent and accurate language) than tasks which are based on abstract and remote information. (see also Robinson, in press). Bygate (1996) showed that familiarity with a task's contents has an effect upon a learners' syntax and lexis. Repeating a cartoon narrative after a three day gap resulted in the subject producing more complex and precise language than he had used to tell the story the first time around. Skehan and Foster (1997b) also researched story-telling tasks and report that a narrative containing a well structured and obvious storyline resulted in significantly more fluent language (measured by false starts, repetitions and reformulations) than a narrative with a less structured storyline. Other research has looked at the way different task goals can affect the language produced to transact them. Brown (1991) concludes that tasks requiring learners to interpret information lead to greater complexity in language and a willingness to hypothesise. Skehan and Foster (1997a) report that a decision-making task involving tricky dilemmas of one sort or another are likely to lead to more complex (i.e. more subordinated) language than a problem-solving task where the available solutions are straightforward and more circumscribed. Robinson, Ting and Urwin (1996) conclude that a task set in the 'here and now' imposes a lighter cognitive load than one set in the 'there and then' as it produced significantly more fluent language.

Taken together, these studies and others like them suggest that the performance of second language learners in terms of their fluency, accuracy and complexity is sensitive to the cognitive demands of tasks, and therefore the cognitive dimension of task design is a very important consideration. From the complex pattern of the evidence it is clear, however, that an assessment of the cognitive demands of particular tasks is not a matter of a simple taxonomy.

5.5 Summing up

This chapter has looked at ways a language learner's attention might be manipulated by classroom choices. The focus-on-form approach seeks to maintain a primary focus on meaning with shifts of attention to form deftly organised by the teacher at moments deemed opportune and on linguistic forms deemed ripe for attention. A different approach uses pre-task planning time to ease the cognitive load imposed by the task so that attentional resources are freed up which may then be available to the learner for improved language performance. Another consideration, and one not sufficiently accounted for in these first two approaches, is the way task characteristics themselves may cause attention to be diverted by the task content or task procedures, or the way task characteristics may predispose learners to prioritise either complexity, accuracy or fluency in their task performance. The focus-on-form approach appears to be of less interest than the other two because of its restricted role within Communicative Language Teaching. It is not easy to see how selective and focussed recasting, for example, could be sufficient treatment for L2 errors outside a content-based syllabus. (As noted above, it is not a straightforward matter to target and recast complex syntactic problems.) The other approaches, those of using pre-task planning or task design to influence where learners are able to direct attention, seem more promising though heretofore they have been researched separately. The research study reported in the following chapters of this thesis is an attempt to investigate the interaction of these two influences on the language produced during tasks, and to see what outcomes might be beneficial to the learning process.

Chapter Six

The Research Study

The preceding chapters have considered language from an information-processing perspective. A central point has been that attention is not an unlimited resource, and in situations where we need to attend to several things at once we have to choose which areas to prioritise. Using language is a complex task requiring simultaneous processing at many different levels. For this to be possible the brain must be able to use a considerable degree of parallel, automatic (i.e. capacity-free) processing. If the processing of, for example, lexis, phonology and syntax could not work in parallel and was instead serial and capacity-demanding, 'speaking would be more like playing chess: an overt move now and then, but mostly silent processing'. (Levelt 1989: 27).

The processing of language is not, however, entirely capacity-free. If it were, it should be possible to carry on a conversation while attending to another capacity-demanding task, such as working out a maths problem. Sadly this is not the case. But even without the distraction of another task, language appears to vary in the amount of attentional capacity it requires. Research into native-speaker language production, reviewed in chapter two above, suggested but did not convincingly prove, that unplanned speech is characterised by a low degree of syntactic complexity because complex language requires more attention than extemporaneous execution can allow. Nevertheless, it is surely the experience of every language user that complex subject matter can make speaking or reading slow and effortful, as can complex syntax.

For many second language learners, perhaps especially those who are just beginning to study the target language, speaking is somewhat like playing chess, with lots of silent

processing interrupting the flow of utterances. Progress in learning a language is characterised by increasing speed of production and comprehension as the amount of attention the brain needs to give to lexical, syntactic or phonological processing is reduced through automatisation. The research reviewed in chapter five above has indicated that planning time enables learners to produce more complex and varied language than they are capable of when they have no time to plan, because of the greater attention planning time affords to these dimensions of their language

From an information-processing point of view, another factor competing for a language-user's attentional resources is the subject matter of the discourse, though this has largely been ignored in the research into the effects of planning time on both the native and non-native speaker language. Because attentional resources are limited, it is likely that complex subject matter will divert attention away from language form (cf VanPatten 1990 discussed in 3.4). For native speakers this might be manifested in dysfluency, and/or a greater reliance on simple syntax. For second language learners, this might be manifested in dysfluency, syntactic simplicity and/or an increase in errors.

The recent interest in the ways language performance might vary according to the amount of attentional resources available has restricted itself to the effect of planning time on the linguistic output of non-native speakers in a task-based context (Ellis 1987, Crookes 1989, Ortega 1995a). The study reported in this thesis draws upon this research, but seeks to widen the field of interest in two ways: firstly to include native speakers as well as non-native speakers, and secondly to take account of the cognitive difficulty of the subject matter of the tasks used in the investigation. Unlike other studies in the field therefore, this study will be able to explore the relationship between planning time (given or denied), speaker type (native or non-native), and task demands (easy or difficult). In a further contrast to previous studies which focus only on the relationship between attentional resources and syntax, this research will look closely at the way way both native and non-native speakers exploit lexis rather than syntax as a way of overcoming insufficient attentional capacity (Pawley and Syder 1983, Ellis 1996, see section 2.4 above)

In the sections that follow, the research hypotheses will be set out and the research methods described in detail. This will include a description of the tasks used, the planning conditions, the subjects and setting, and the research design.

6.1 Hypotheses

The overriding hypothesis of this research is that greater attentional resources, operationalised by giving pre-task planning time, will have positive effects on language performance in terms of increased fluency, greater syntactic complexity and, in the case of the non-native speakers, increased accuracy. This general hypothesis can be broken down into several more detailed hypotheses.

Hypothesis 1. If planning time is given before a task, less on-line processing will be required to organise ideas. Under planned conditions therefore, there will be greater fluency in language performance: that is, there will fewer repetitions, fewer reformulations, fewer replacements, fewer false starts, fewer pauses and a smaller proportion of silence to speech.

Hypothesis 2. If planning time is given before a task, a speaker can organise more complex ideas and be more ambitious in expressing them. Under planned conditions therefore, there will be greater subordination in the language: that is, there will be a greater number of clauses per AS[23] unit. Additionally, and following Ochs (1979), planning time will be associated with a more syntactic than pragmatic mode of speech, and will result in an increased incidence of verbs in the passive voice.

Hypothesis 3. If planning time is given before a task, a speaker will rely less on pre-formed lexicalised language to express ideas. It is reasoned that the use of such lexicalised

---

[23] Analysis of Speech unit, as defined by Foster, Tonkyn and Wigglesworth (forthcoming). This unit of analysis is based on the c-unit and T-unit used by other researchers in the field, but has a much more detailed definition and allows for a principled analysis of nearly all the elliptical fragments so pervasive in oral language.See Appendix D

language is a time-gaining processing strategy that will be most evident when attentional resources are scarce. Under planned conditions therefore, there will be fewer lexicalised 'chunks' in the language produced.

Hypothesis 4. It is reasoned that for native speakers the processing of syntactic and morphological rules is largely automatised and therefore little affected by the amount of attention available. For non-native speakers only then, if planning time is given before a task there will be greater accuracy in the language produced. This hypothesis is motivated by the reasoning that greater accuracy should result from the greater attention made available on-line to speakers who have planned what they will say. The lack of such as effect in the Crookes (1989) study might be accounted for by the narrow and specific accuracy measures used. Using a more global measure of accuracy therefore, under planned conditions there will be a higher proportion of error-free clauses per total clauses.

Hypothesis 5. For non-native speakers only (because this condition will not be applied to the native speakers) the effects predicted in hypotheses 1-4 will be greater when the planning is not solitary but guided by suggestions on how to proceed. It is proposed that guided planning will lead subjects to use their time more efficiently and allow them to plan more effectively.

Hypothesis 6. The effects predicted in hypotheses 1-5 will be greater for tasks that are more cognitively demanding, i.e. those that require more on-line processing. It is reasoned that more cognitively demanding tasks require more of a speaker's attentional resources and therefore provide greater scope for pre-task planning.

Hypothesis 7. There is no existing research that compares the task-based performance of native and non-native speakers under planned and unplanned conditions. Therefore this study will adopt the null hypothesis. There will be no difference in the way planning time affects the fluency, complexity or lexical choices in the performance of native and non-native speakers.

6.2 Tasks

All data was recorded from subjects working on three tasks with the same partner each time. The three tasks used were based on type commonly found in language textbooks[24], and their contents were reasoned to require different levels of attention from the subjects, with progressively less familiar and less predictable information causing an increasingly taxing cognitive load.

The **personal information exchange** task required subjects to describe to the other member of the dyad how to get to his or her home from the college (where the research was carried out) in order to turn off a gas cooker that had been left on. As it involved accessing information well known to the speaker and possibly already rehearsed in real-life, it was seen as requiring the least cognitive effort and therefore allowing the greatest attention to language form. Moreover, it was reasoned that the nature of the task would require relatively simple linguistic forms, relatively undemanding therefore of cognitive attention. (See Appendix A)

For the **narrative task**, each member of the dyad had to construct a storyline from a set of five pictures that were loosely but not obviously connected, and to relay their ideas to each other. This task involved encoding new, visual information into linguistic form, and required some degree of imagination. It was seen as giving scope for more complex language but also demanding greater cognitive effort, therefore allowing less attention to be devoted to language form.

For the **decision-making** task, subjects were asked to act as judges at the trials of a list of offenders and to reach an agreement with their partner on a suitable prison term for each. This task involved considering a lot of new information (i.e. the facts of each case), evaluating it, and then defending an opinion against any objections from the other side of the dyad. This task was reasoned to place the heaviest cognitive load upon the subject and to allow the least attentional resources to be given to language from. At the same time,

---

[24] A number of candidate tasks were piloted at a neighbouring college in order to determine for each of the three types which task was the most productive of language and the most enjoyable to do. In this way, the risk of basing the research project on unpopular and unproductive tasks was reduced.

the process of trying to reach an agreement on a series of difficult questions was considered most likely to require the use of complex language. (See Appendix C)

Summing up the characteristics of these three tasks, the personal information exchange is judged to be the easiest because it involves well-known, accessible information and can be transacted in relatively simple language. The decision-making task is judged to be the most difficult because it involves coping with a lot of new information, and formulating possibly complex arguments. At the same time, this task is the only one of the three which is interactional in design and which therefore has unpredictable outcomes. The narrative is judged to be more complex than the personal task because it requires the processing of new information, but less complex than the decision-making task because it does not involve unpredictable interactions with a partner.

6.3 The subjects and setting

Initially sixty-four subjects took part in this study. Thirty-two were native speakers of English, enrolled as first-year undergraduates at St. Mary's University College in Twickenham. All were aged between 18 and 20, and only five were male.They were majoring in a variety of fields (sports science, geography, history, religious studies, drama, sociology and English literature) and were all in the same two class groupings (class NS 'A' and class NS 'B') for a required English Language first year core course concentrating on the role of accent and dialect in society. None of them had ever done any previous study of the English Language.

Data gathering took place during normal class time. Subjects were asked if they would be willing to be recorded while carrying out three kinds of language tasks used in second language classrooms. All the students happily agreed, though it was made clear that anyone unwilling to take part was free to leave. Logistically, it was not possible to assign the students randomly to either the planned or unplanned condition. Therefore all the students in one class grouping (NS 'A') were given the tasks without planning time, and

all who were in the other class grouping (NS 'B') were given the tasks with planning time. When the data was transcribed it was found that one male student in the planned condition had failed to carry out the task as set, and was therefore excluded from the analysis, giving a native speaker subject total of thirty-one.

There were also thirty-two non-native speaker subjects. They were pre-intermediate level students studying English as a Foreign Language 6 hours per week at Richmond Adult and Community College centres in Twickenham and Richmond. They came from wide variety of L1 backgrounds and were all between 18 and 30 years old. Only three were male. They had been placed in one of four pre-intermediate level classes on the basis of a brief interview and the college placement test. Which of the four classes (NNS 'A', 'B', 'C' or 'D') an individual student chose to join was not based on any difference in proficiency, but on a personal preference for college venue and time of day. Eight subjects were chosen for study from each of the four classes, although data was collected from everyone attending. The students selected for study were simply those who were present at each of the three data gathering sessions and who worked with the same partner each time.[25] When the recordings were analysed it was found that one male subject was of a markedly lower proficiency than the others. It was subsequently revealed that, unusually, he had not taken the placement test and was in a pre-intermediate class for reasons of scheduling convenience. He was excluded from the analysis, giving a total of thirty-one non-native speaker subjects. The total number of subjects in this study is therefore sixty-two.

6.4 Procedure

All data was collected during scheduled class times. The researcher, who was familiar to all the native speakers and to many of the non-native speakers, took the role of the teacher and introduced the tasks. This was to ensure that the implementation was identical each

[25] There was considerable attrition in the non-native speaker grouping, due to absences over the three weeks of data-gathering. This had been anticipated, however, and the remaining thirty-two subjects, out of a possible pool of nearly 100, was considered to be quite adequate for the research purposes. The subject

time. The tasks themselves were presented to the non-native speakers as straightforward classroom communication activities (which indeed they were, having been modelled on those found in typical language textbooks). To the native speakers the tasks were presented as second language classroom activities for which native speaker baseline data was sought. None of the participants was aware of the difference in implementation between the planned and unplanned conditions, nor of the research hypotheses.

All data was collected using small, unobtrusive dictation machines with no external microphone, which enabled all the dyads from a particular class grouping to be recorded at the same time in the same room. It is argued that the use of intact classes under relatively normal conditions minimises any effects that experimental conditions might have on the subjects performance, especially perhaps in the case of the non-native speakers who could be inhibited or intimidated by being recorded alone in front of large equipment (Foster 1998). In addition, the inclusion of all class members in the data gathering meant that the subjects selected for analysis (by regularity of attendance) were not alerted to or troubled by the fact that only their data would be analysed.

6.5 Planning condition

The control groups, one native speaker and two non-native speaker, carried out the three tasks with only a brief introduction to ensure that they understood what the task required of them. (see Appendix C). The three experimental groups, one native speaker and two non-native speaker, were given the same brief introduction to the task and then 10 minutes of individual planning time during which they were expected to make notes about what they would say, although these notes would be taken away when the task proper started. For the two non-native experimental conditions only, the planning condition was operationalised as detailed and undetailed. The subjects in the undetailed planning condition were treated in exactly the same way as those subjects in the native speaker

---

pool of the native speakers was much smaller, only forty, but data for the three tasks was collected at the same session, thus avoiding the danger of disastrous attrition through absences.

planning condition: they were merely told to plan what they would say in the task. (See Appendix A) The non-native subjects who were in the detailed planning condition were given written guidance on how they might best use the 10 minutes planning time to consider the lexis, syntax content and organisation of what they would say. In this way it was hoped that it would be possible to see what effect quality of planning might have on non-native speaker performance. (See Appendix C.)

6.6 Design

The rather complex research design is illustrated in table 6.1 below.

Table 6.1: Design of the Control and Experimental Groups

(NS = Native speaker; NNS = Non-native speaker)

Control Groups

with no pre-task planning time
(total of 32 subjects)

NS 'A'	NNS 'A'	NNS 'B'
16 subjects	8 subjects	8 subjects

Experimental Groups

with ten minutes pre-task planning time

(total of 32 subjects)

NS 'B'	NNS 'C'		NNS 'D'	
(undetailed planning.)	(undetailed planning.)	(detailed planning.)	(undetailed planning.)	(detailed planning.)
16 subjects	4 subjects	4 subjects	4 subjects	4 subjects

The study is essentially a multifactorial design, examining three task types, two speaker types and two implementation conditions. Half of the native speakers (those in class NS 'A') were assigned to the unplanned (control) condition, and the other half (those in class NS 'B') were assigned to the planned (experimental) condition. In a similar fashion, half of the non-native speakers (those in NNS 'A' and NNS 'B') were assigned to the unplanned (control) condition, and the other half (those in NNS 'C' and NNS 'D') were assigned to the planned (experimental) condition. For the non-native speakers only it was possible to further divide the planned condition into detailed and undetailed planning as described above in 6.4. In order to do this, each of the non-native subjects in the planned condition was randomly assigned to either an undetailed or detailed planning condition. As all planners, whether detailed or undetailed, were given 10 minutes of individual planning time, it was not obvious to any of them they might have a different (i.e. more or less detailed) set of instructions from others in the group. When the tasks began, subjects worked with a partner who had received exactly the same pre-task instructions.

In the case of the non-native speakers, the tasks were administered at weekly intervals. To combat any practice effect, each of the controls and each of the experimental groups did the tasks in a different order. The eventual scores for the Narrative task were therefore made up of subjects who had done this task in week one of the study and those who had done it in week three. The scores for the Decision task were similarly spread, rendering the two condition directly comparable. It was assumed that, if there was a practice effect of consistent strength over the three visits of the study, the Personal task, always done in the second week, would be affected to the same degree as the averaged performance on the other two tasks.

In the case of the native speaker subjects, it was felt that the students would certainly see each other every day on the campus and would very likely talk about the tasks, (something that would not happen to the non-native speakers who were all part-time and who did not see other students from other classes). Therefore it was decided to administer all three tasks to all the native speaker subjects on the same day to avoid the danger of tasks being

discussed and prepared in advance. The practice effect was countered by half the control group and half the experimental group doing the Narrative task first, while the other half did the Decision task first. All native speaker subjects did the Personal task second, and whichever task remained was done last. Thus the native speaker scores for all the tasks were averaged out in the same way as described above for the non-native speakers. This task ordering is set out in table 6.2 below.

Table 6.2: Task order across the groups

(NS = Native speaker: NNS = Non-native speaker.)

	Control Groups			**Experimental Groups**		
	NS 'A'	NNS 'A'	NNS 'B'	NS 'B'	NNS 'C'	NNS 'D'
Time 1	Decision or Narrative	Decision	Narrative	Decision or Narrative	Narrative	Decision
Time 2	Personal	Personal	Personal	Personal	Personal	Personal
Time 3	Narrative or Decision	Narrative	Decision	Narrative or Decision	Decision	Narrative

In the non-native speaker condition, all students present at the class were recorded, including those whose previous absences had meant they would not be part of the study. In this way, subjects were protected from any pressure that selective recording might have had upon their performance. At the end of the three weeks, data was transcribed only for the subjects who had attended each of the three sessions and worked with the same partner each time, giving a total of 32 subjects. (Later, the student of very low proficiency was identified and excluded, bringing the total down to 31). One subject in the control group failed to take her turn in one of the tasks, and an experimental dyad who had only attended for two of the sessions was nevertheless included in order to even up the numbers across the four groups. The data set, based on 31 subjects, is otherwise complete. The same procedures were followed for the native speakers, with the exception,

mentioned above, of all data being gathered on the same occasion. The data set for 31 native speaker subjects excludes the one subject who failed to complete the tasks as set but is otherwise complete.

6.6 Measures used in the analysis

The recorded data was transcribed and various measures were used to analyse the fluency, accuracy and complexity of the language produced during the tasks. The fluency measures chosen were those of reformulation, replacement, false-starting, repetition and pausing. They were used to assess two different aspects of fluency. Following Skehan (1998), the first three (reformulation, replacement and false-starting) are reasoned to indicate the frequency of breakdown repair needed to maintain continuous speech. The other two (repetition and pausing) are reasoned to measure the amount of on-line processing involved in the production of speech. Pausing is considered significant both in its frequency and its length. These fluency measures are defined as follows:

Reformulations: Either clauses or sub-clausal items that are repeated with some modification to syntax, morphology or word order.

Replacements: Lexical items that are immediately substituted for another.

False starts: Utterances that are abandoned before completion. They may or may not be followed by a reformulation.

Repetitions: Words, sub-clausal items or clauses that are repeated with no modification whatever to syntax, morphology or word order.

Pauses: A break of 1.0 second or longer either within or between turns.[26]

Silence total: The number of seconds silence represented by the sum of the pauses in each transcript.

Hypothesis three deals with another aspect of fluency: the use of pre-formed lexicalised 'chunks' of language to reduce the burden of syntactic processing. The definition of what

---

[26] According to Jefferson (1997) pauses of less than 1.0 second are not noticed or marked in fluent conversation.

constitutes such a chunk is problematic, and consequently so is the identification of such chunks in the data. Because of the special difficulties such an analysis presents it will be dealt with separately and at length in chapter eight below.

Two measures are used to assess complexity in language. Firstly, and more simply, the number of verbs in the passive voice is taken as an indication of whether the mode of speech is syntactic (i.e. relatively complex) or pragmatic (i.e. relatively simple) as described by Ochs (1979).

Secondly, syntactic complexity is measured by calculating the number of subordinate clauses (defined below) per basic speech unit. There are many ways of dividing language into basic units, either syntactically, semantically or intonationally, but little agreement on either which is the most valid or on how these units should be defined. (See Foster, Tonkyn and Wigglesworth (forthcoming) for a full discussion of this problem). This research study uses the AS-unit, level two analysis, as defined by Foster et al. (forthcoming). The AS-unit is a development of the more well-known T-unit and c-unit, i.e. it is syntactically based, but more extensively defined in order to cope with the many elliptical and fragmentary elements of oral interaction that classic T-unit or c-unit analysis has to exclude. (See Appendix D for greater detail.)

Briefly, the AS-unit can be defined as follows:

AS–unit: An utterance consisting of either an independent simple clause, or sub-clausal unit, together with any subordinate clause(s) associated with either. An independent simple clause will be minimally a clause including a finite verb.

That's right, turn left (two AS-units)
I take a different way (one AS-unit)
You go to the main street of Twickenham (one AS-unit)

An independent sub-clausal unit will consist of one or more phrases which can be elaborated to a full clause by means of recovery of ellipted elements from the context of the discourse or situation.

A: which time the delivery of my home (one AS-unit)
B: about er maybe 9 a.m. (one AS-unit)

A: Oh poor woman! (one AS-unit)

Highly interactional data typically yields a considerable proportion of minimal units (e.g. one-word minor utterances and echoic responses) whose inclusion in an analysis could distort the perception of the nature of the performance. Thus, in a level two AS-unit analysis the following one-word minor utterances are excluded:

Yes; No; Okay; Uhuh; Right

Also excluded are verbatim echo responses, as in B's utterance below:

A: I think two years
B: Two years

In order to measure complexity of language through subordination it is necessary to identify subordinate clauses within the basic syntactic unit. Subordinate clauses and complexity in language are defined thus:

Subordinate clause: Minimally this consists of a finite or non finite verb element plus at least one other clause element (Subject, Object, Complement or Adverbial).

I have no opportunity to visit. (1 AS-unit, 1 clause)
And you you be surprise <u>how he can work</u> (1 AS-unit, 2 clauses)

Complexity of language: This is calculated as the total number of clauses, both independent and subordinate, divided by the total number of AS-units.

Accuracy in language can be measured specifically (as in Crookes 1989, Ellis 1987, Ortega 1995a) or more globally by looking at the proportion of errors overall. For these research purposes it was considered that a global measure would be better at capturing possible variances in accuracy across the conditions. This global measure is sometimes expressed as the proportion of T-units or c-units that contain no errors, but as these units can be quite extended, and therefore less likely to be completed without some error occurring, the scores achieved can be very low. The following example demonstrates this:

A: I thinks I need to go to er ask the manager
B: oh well yes
A: because you haven't got a receipt.

There are no error-free T-units or c-units here to credit to speaker A, but there are two error-free clauses. Accordingly this more reliable measure was used to assess accuracy in the data. The precise definitions of an error-free clause and an accuracy score is given below:

Error-free clauses: These are clauses, independent and subordinate, which contain no error in syntax, morphology or word order. Errors in lexis are included if the word used is incontrovertibly wrong, but in cases of fine distinctions of appropriacy no error is recorded.

Accuracy: This is the number of error-free clauses, independent and subordinate, expressed as a percentage of total clauses.

All the data was collected according to the procedures described above. The tapes were transcribed and then coded for each of the dependent variables. The results for each of the hypotheses will be described in the following chapter.

Chapter Seven

Results and Discussion

The non-native speaker data was collected and analysed with Peter Skehan as part of a programme of six collaborative research studies into task-based learning. It was subsequently published as Foster and Skehan (1996). This chapter presents and extends work published in that article. The native speaker data was not part of the collaboration, but was conceived to extend many of the same research questions into the native speaker domain. In the discussions that follow in this and subsequent chapters, the non-native speaker results will usually be referred to as part of the present study, but in places where comparisons are being drawn with other studies into planning and non-native speaker performance, reference will be made to Foster and Skehan (1996).

Statistical analyses of all the data were carried out using SPSSPC software. The procedure for the native speaker data, in which only two planning conditions were being compared, was to carry out a comparison of means through a t-test on each of the three tasks. For the non-native speaker data, in which three planning conditions were being compared, one-way ANOVAS were carried out for each dependent variable for each task, in addition to t-tests.

The results will be discussed in relation to six of the seven research hypotheses. The measurement and analysis of lexicalised language in the research data, with which Hypothesis 3 is concerned, was particularly detailed and complex and so is dealt with separately in chapter eight below. For Hypotheses 1 and 2, the results will be presented first for the native speaker data and then for the non-native speaker data. Hypotheses 4 and 5, which are concerned only with the non-native speaker data, will then be considered. The results for Hypotheses 6 and 7, which are concerned with both the non-native and native speaker data, will be followed by a discussion of the results as a whole.

7.1 Native speaker results for Hypothesis 1

Hypothesis 1 stated that planning time would be associated with greater fluency in performance on the three tasks as evidenced by fewer repetitions, fewer reformulations, fewer replacements, fewer false starts, fewer pauses and a smaller proportion of silence to speech. The results for the native speaker data are presented in Table 7.1.

Table 7.1: Mean scores for the dysfluency measures in the native speaker data

	repetitions		
	personal task	narrative task	decision task
no planning	2.36	3.62	5.69
With planning	1.73	2.5	3.6
	p = .43ns	*p = .25ns*	*p = .28 ns*

	reformulations		
	personal task	narrative task	decision task
no planning	1.0	1.88	2.69
With planning	1.2	1.17	1.73
	p = .65ns	*p = .33ns*	*p = .31ns*

	replacements		
	personal task	narrative task	decision task
no planning	1.21	1.44	0.88
With planning	0.73	.050	.053
	p = .36 ns	*p = .04*	*p = .34 ns*

	false starts		
	personal task	narrative task	decision task
no planning	2.86	5.63	8.0
with planning	2.80	2.59	4.07
	p = .95 ns	*p = .04*	*p = .02*

	pauses		
	personal task	narrative task	decision task
no planning	5.64	20.25	28.25
with planning	3.73	5.67	14.80
	p = .09ns	*p < .000*	*p < .000*

	total of silence		
	personal task	narrative task	decision task
no planning	7.93 secs	40.88 secs	41.38 secs
with planning	4.67 secs	7.17 secs	20.0 secs
	p = .062ns	*p = .001*	*p < .000*

Repetitions were defined as words, sub-clausal items or clauses that are repeated without modification to syntax, morphology or word order. The results show a general trend across all three tasks for planning time to reduce the number of repetitions, as predicted by Hypothesis 1, but none of the scores reaches statistical significance ($p = 0.43$, $p = 0.25$, $p = 0.28$). For this measure therefore, Hypothesis 1 is not confirmed.

Reformulations were defined as clauses or sub-clausal items that are repeated with some modification to syntax, morphology or word order. The pattern of results here is not consistent across the tasks. For the decision and narrative tasks the planning condition reduces the mean totals for reformulations (from 2.69 to 1.73, and from 1.88 to 1.17) Although this is in the direction predicted by Hypothesis 1, the figures are not statistically significant ($p = 0.31$ and 0.33). For the personal task the planning condition slightly increases the mean total of reformulations from 1.0 to 1.2., i.e. in the opposite direction to that predicted by Hypothesis 1, but this is very far from significance ($p = 0.65$). For this measure therefore, Hypothesis 1 is not confirmed.

Replacements were defined as lexical items that are immediately substituted for another. The results indicate that whatever the task or planning condition, replacements occurred in the data infrequently. All the mean scores for the unplanned condition are between 0.88 and 1.44, and for the planned condition are between .05 and 0.73. Although there is a consistent trend across all tasks for planning time to reduce the number of replacements, this does not reach statistical significance in the personal and decision tasks ($p = 0.36$, $p = 0.34$). It is only in the narrative task that a significant effect for planning is shown, with a mean total of replacements for the unplanned condition of 1.44 compared to a mean total

for the planned condition of .05 (p = 0.04) For this measure therefore, Hypothesis 1 receives partial confirmation.

False starts were defined as utterances abandoned before completion. For the personal task, the mean totals are almost the same (2.86 for the unplanned condition and 2.80 for the planned condition, p = 0.95) For the other two tasks, the mean totals reveal that false starts were significantly fewer in the planned condition. In the narrative task, the mean total is 5.63 for the unplanned condition and 2.59 for the planned condition (p = 0.04). In the decision task the mean total is 8.0 for the unplanned condition and 4.07 for the planned condition (p = 0.02). For this measure, Hypothesis 1 receives partial confirmation.

The planning condition results in fewer pauses of greater than one second for all three tasks. For the personal task, this difference is small and not significant (5.64 compared to 3.73, p = 0.09). For the other two tasks the difference is much greater and reaches high levels of significance. For the narrative, in the unplanned condition the mean total of pauses is 20.25 compared to only 5.67 in the planned condition. For the decision task, in the unplanned condition the mean total of pauses is 28.25 compared to 14.80 in the planned condition. Both these results achieve a level of significance in which p = .000. There is thus strong confirmation for Hypothesis 1 for the narrative and decision tasks, but none for the personal task.

The measure of dysfluency to be considered is that of the proportion of silence to speech, calculated by adding up all the pauses of greater than one second. The figures reveal that the planning condition results in a smaller proportion of silence to speech across all three tasks, as predicted by Hypothesis 1. For the personal task, the totals are 7.93 seconds in the unplanned condition compared to 4.67 seconds in the planned condition, a difference which just fails to achieve significance (p = .062). In the other two tasks the differences are greater. For the narrative, in the unplanned condition the mean total of silence is 40.88 seconds compared to only 7.17 seconds in the planned condition (p = .001). For the decision task, in the unplanned condition the mean total of silence is 41.38 seconds

compared to 20.0 seconds in the planned condition ($p = .000$) Thus Hypothesis 1 receives strong confirmation for the narrative and decision, but is not confirmed for the personal task.

We can sum up these native speaker fluency results by noting that, with the one exception of reformulations in the personal task, the trend for all measures and all tasks is for planning time to be associated with greater fluency. However, this trend is generally not strong enough to reach statistical significance except in the case of number of false starts, pauses and proportion of silence in the narrative and decision tasks, and the number of replacements in the narrative task. It is notable that the trend for greater fluency is always weakest in the personal task, with none of the measures achieving statistical significance. This is in accordance with Hypothesis 6 which predicts that the more cognitively demanding narrative and decision tasks will show stronger effects for planning than the cognitively less demanding personal task. This observation will be expanded on below in section 7.7.

7.2 Non-native speaker results for Hypothesis 1

We turn now to the non-native speaker fluency scores. Table 7.2 below gives the separate *F* values for one-way ANOVAs relating the three levels of planning (none, undetailed and detailed) to the each of the fluency measures for each of the three tasks

Table 7.2: *F* values for fluency and task type in the non-native speaker data .

	personal	narrative	decision
replacements	1.59	0.37	4.48*
false starts	0.15	0.13	0.84
reformulations	1.28	0.62	0.45
repetitions	0.32	0.79	2.99*
pauses	4.68*	9.3***	12.33***
total of silence	3.47*	17.31***	21.94***

** $p<0.5$ *** $p<.001$*

The F values for number of pauses and total of silence reach significance for all three tasks, with very high levels achieved for the narrative and decision tasks. The other significant F values are for replacements and repetitions in the decision task alone. Accordingly, the mean scores for only these variables are given in Table 7.3 below.

Table 7.3: Mean scores for the dysfluency measures in the non-native speaker data.

	number of pauses		
	personal task	Narrative task	decision task
no planning	19.2	30.3	37.0
undetail. planning	10.0	15.3	17.3
detailed planning	11.2	8.0	17.3

	total of silence		
	personal task	Narrative task	decision task
no planning	31.8 secs	120.3 secs	91.4 secs
undetail.planning	19.5 secs	29.3 secs	25.5 secs
detailed planning	14.5 secs	14.2 secs	29.5 secs

	Decision task	
	Repetitions	replacements
no planning	9.9	.056
undetailed planning	20.1	1.38
detailed planning	17.9	2.63

For all three tasks there are significant effects for planning on the number and length of pauses. In the no planning condition, the mean total for number of pauses is 19.2. This compares significantly ($p < .05$) with the undetailed planning condition result of 10.0 and the detailed planning condition result of 11.2. The effects of planning on the narrative task are more significant ($p < .001$) with the no planning condition resulting in a mean total for number of pauses of 30.3 compared to 15.3 in the undetailed planning condition and only 8.0 in the detailed planning condition. Similarly in the decision task, the no planning condition produced a mean total number of pauses of 37.0 compared to 17.3 for both the detailed and undetailed planning condition. It is interesting to note that the cognitively

more demanding narrative and decision tasks are associated with much more frequent pausing than the less taxing personal task, but that the difference between the mean totals is diminished by planning time. We can conclude that for the measure of number of pauses Hypothesis 1 is confirmed for all three tasks

There are significant effects also for planning on the proportion of silence to speech in all three tasks. Once again it is the personal task where the effect, though significant, is not as strong as on the other two tasks. The mean total of silence for the personal task is 31.8 seconds, which compares significantly ($p < .05$) with 19.5 seconds for the undetailed planning condition and 14.5 seconds for the detailed planning condition. The impact of planning time on the total of silence in the narrative and decision tasks is much greater. In the no planning condition the mean total for silence in the narrative tasks is a remarkable 120.3 seconds, comparing very significantly ($p < .001$) with a mean total of only 29.3 seconds in the undetailed planning condition, and only 14.2 seconds on the detailed planning condition. A similar effect is apparent in the decision task where a mean total of 91 seconds silence in the no planning condition compares very significantly ($p < .001$) with the mean total of 25.8 seconds in the undetailed planning condition and 29.5 seconds in the detailed planning condition. For the proportion of speech to silence, these results provide strong confirmation for Hypothesis 1.

The *F* values for the one-way ANOVAs did not provide confirmation for Hypothesis 1 on any other fluency measure. Interestingly, the two significant *F* values found for repetition and replacement in the decision task only were actually in the opposite direction to that predicted by Hypothesis 1. That is to say that planning was associated with *more* repetitions and replacements in the decision task rather than with fewer. The mean total for repetitions in the no planning condition is 9.9. This rises to 20.1 in the undetailed planning condition and to 17.9 in the detailed planning condition ($p < .05$). For replacements, the no planning condition has a mean total of only .056. This rises to 1.38 in the undetailed planning condition and 2.63 in the detailed planning condition ($p < .05$). The opposite is true in the native speaker data. Although the native speaker results only

achieve significance once in these measures (for replacements in the narrative task) the general trend is for planning to be associated with fewer incidences of replacements and repetitions in all three tasks. This complex pattern will be discussed further in section 7.10.2.

7.3 Native speaker results for Hypothesis 2.

Hypothesis 2 predicted that planning time would result in language that was more complex, i.e. that there would be a greater number of clauses per AS-unit, and a higher incidence of the passive voice. The results for the native speaker data are given below in Table 7.4.

Table 7.4: Mean scores for complexity in the native speaker data

	clauses per AS-unit		
	personal task	Narrative task	decision task
no planning	1.20	1.24	1.36
with planning	1.30	1.62	1.76
	p = .09ns	*p = .002*	*p = .000*

	Passives		
	personal task	Narrative task	decision task
no planning	0	1.75	2.56
with planning	0.13	3.25	1.67
		p = .114	*p = .095*

These results show that for all three tasks planning is associated with more subordinated language. For the planning condition in the personal task this increase does not reach significance, (p = .09) but there is a strong effect for planning on the other two tasks. In the narrative, a mean total of 1.24 clauses per AS-unit in the no planning condition compares very significantly (p = .002) with 1.62 in the planning condition. The comparison in the decision task of 1.36 clauses per AS-unit in the no planning condition with 1.76 in the planning condition is even more significant (p = .000). However, the results for the incidence of the passive are not so straightforward. There were too few in the personal task data for any meaningful analysis to be made. Passives also occurred

rather infrequently in the other two tasks, and planning time had an opposite effect in these, producing more in the narrative and fewer in the decision task (with neither score achieving significance). Hypothesis 2 is therefore supported for the measure of subordination in the narrative and decision tasks, though not in the personal task, and is not supported at all for the incidence of passives.

7.4 Non-native speaker results for Hypothesis 2

Following the sequence of analyses used for the fluency measures in the non-native speaker data, Table 7.5 below shows first the *F* values for the one way ANOVAs on the number of clauses per AS unit, and then the mean scores for the three planning conditions.

Table 7.5: *F* values and mean scores for complexity in the non-native speaker data

	Passives		
	personal task	Narrative task	decision task
F values	0.49ns	2.38ns	1.2ns
	clauses per AS unit		
	personal task	Narrative task	decision task
F values	8.35***	9.30***	15.00***
mean scores			
no planning	1.11***	1.20***	1.23***
undetail.planning	1.16***	1.43***	1.35***
detailed planning	1.26***	1.68***	1.52***

*** $p < .001$

The results in Table 7.5 show no significant effect for planning on the incidence of passives on any of the tasks, but a strong effect for planning on subordination across all tasks ($p < .001$). In each case the detailed planning condition produced significantly more subordinated language than the undetailed planning condition, which in turn produced significantly more subordinated language than the no planning condition. Range tests using Duncan's procedure (Norusis 1990) indicate that within each task all pairwise comparisons within each task are significantly different from one another with the

exception of the no planning (1.11) and undetailed planning (1.16) conditions for the personal task.

Hypothesis 2 is accordingly not supported for passives but upheld for the measure of subordination. This pattern of results showing a greater effect on the degree of subordination for detailed than undetailed planning is also support for Hypothesis 5 and will be considered below in section 7.6

7.5. Non-native speaker results for Hypothesis 4

Hypothesis 4 predicted that planning time would be associated with greater accuracy in the language produced by the non-native speakers. Accuracy was measured by calculating the number of error-free clauses as a percentage of total clauses. Table 7.6 below shows the three *F* values from the one-way ANOVAs, followed by the means scores.

Table 7.6: *F* values and mean scores for accuracy in the non-native speaker data

	Personal	narrative	Decision
F values	2.46*	.69ns	5.73**
	Percentage of error free clauses		
means scores			
no planning	64%	61%	63%
undetail. planning	76%	65%	73%
detailed planning	69%	58%	71%

*$p < .05$ **$p < .01$

The *F* value for the personal task is significant at the .05 level, and that for the decision task is significant at the .01 level, but significance is not attained for the narrative task in which the detailed planning condition produced results in the *opposite* direction to that proposed by the hypothesis, i..e. the language was less accurate than that in the no planning condition (58% compared to 61%). Duncan range tests demonstrate that the no planning condition was significantly different from the two planning conditions but the two types of planning did not produce significantly different effects. However, the mean scores

for the undetailed planning condition show that the trend for each task is the same and in the direction predicted by the hypothesis, namely that planning time would be associated with greater accuracy in the language produced. The personal and decision task scores for the detailed planning (69% and 71%) are lower than that for the undetailed planning condition (76% and 75%) but unlike the scores in the narrative task, remain above the accuracy levels of the no planning condition (64% and 63%). Hypothesis 4 is therefore supported for the personal and decision tasks but not for the narrative task.

7.6 Results for Hypothesis 5.

This hypothesis concerned the non-native speaker data only and proposed that the effects of planning would be greater when the planning is not solitary but guided by suggestions to the subjects on how to proceed. It predicts that there will be a consistent pattern in the results in which detailed planning will affect language not only in the same way as the undetailed planning, but also to a greater degree. The hypothesis can be explored by considering the non-native speaker results already presented in the tables above.

In the case of the fluency measures of number of pauses and total of silence, the no planning condition consistently showed the least fluency, but the detailed planning condition did not have the predicted effect of producing the greatest fluency except in the narrative task (for both number of pauses and total of silence) and in the personal task (for total of silence only). In the narrative task the undetailed planning condition produced fewer pauses than the no planning condition (a mean total of 15.3 compared to a mean total of 30.3) with the detailed planning condition continuing this trend by producing even fewer (a mean total of 8.0). The results for the total of silence in the narrative follow the same pattern. Here the mean total in the no planning condition was 120.3 seconds. This compares to a mean total of only 29.3 seconds in the undetailed planning condition and the even lower mean total of 14.2 seconds in the detailed planning condition. In the personal task the no planning condition produced a mean total 31.8 seconds silence,

compared to 19.5 seconds in the undetailed planning condition and only 14.5 in the detailed planning condition. For these instances only, Hypothesis 5 is supported.

However, the full picture is more complex. There is one other place in the fluency scores where the detailed planning condition has a significantly stronger effect than the undetailed planning condition, and this is for the mean total of replacements in the decision task. Here the no planning vs undetailed planning vs detailed planning scores are indeed in a monotonic relationship (.056 vs 1.38 vs 2.63) but this is in the opposite direction to that predicted by Hypothesis 5. This is to say, in the case of replacements, undetailed planning produced more dysfluency than no planning, and detailed planning produced the most of all.

In the personal and decision tasks, the effect of the detailed planning condition on the mean totals for pausing and silence is in one instance exactly the same as that of the undetailed planning condition (number of pauses in the decision task), but otherwise the results are in the opposite direction to that predicted by Hypothesis 5. A different pattern is seen in the number of repetitions in the decision task, which is the only other place where the results for the fluency measures achieve statistical significance. Here the undetailed planning condition produced a greater number of repetitions than the no planning condition while the detailed planning condition had a similar but less strong effect (20.1 vs 9.9 vs 17.9).

As far as fluency is concerned therefore, Hypothesis 5 is supported for pausing in the narrative and personal tasks, and for total of silence in the narrative task alone. Otherwise Hypothesis 5 is not supported. For repetitions in the decision task, detailed planning operated in the opposite direction to that predicted and produced a weaker rather than stronger effect than the undetailed planning condition. For replacements in the decision task, the detailed planning condition operated as predicted, i.e. its effects were stronger than that of the undetailed planning condition, but this was in the opposite direction to Hypothesis 1, i.e. planning overall resulted in more rather than less dysfluency. As we have

noted, the other measures of dysfluency failed to reach significance in the ANOVAs, and the mean scores, presented below in table 7.7, show that there is not even a remotely consistent pattern. The detailed planning condition scores move as often in the same direction of the undetailed planning scores as they run counter to them.

For the complexity measure of number of clauses per AS-unit, Hypothesis 5 is strongly supported across all three tasks. The undetailed planning condition always produced significantly more complex language than the no planning condition, with the detailed planning condition producing more complex language than the undetailed planning condition.

The results for accuracy, measured by the percentage of error-free clauses, are never in the predicted monotonic relationship and thus Hypothesis 5 is not supported.. Although the undetailed planning condition always produced more accurate language than the no planning condition, this trend was not continued in the detailed planning condition, but was reversed. For the personal and decision tasks, detailed planning produced only 69% and 71% error-free clauses compared to 76% and 73% in the undetailed planning condition. In the case of the narrative task the detailed planning condition produced less accurate language even than the no planning condition (61% vs 58%). The F value from the one-way ANOVA on the narrative task did not reach significance, in contrast to the other tasks, but the trend in the results is clear: the effect of detailed planning on accuracy is weaker, not stronger, than the effect of undetailed planning.

The results for Hypothesis 5 are mixed, with unambiguous support only in the case of complexity. The hypothesis is not supported at all in the case of accuracy, and only partially in the case of fluency. The finding that the detailed planning condition results in greater complexity but also in more errors and, for some measures, greater dysfluency is interesting, suggesting that it is not easy for non-native speakers to channel extra attention equally towards all aspects of language performance.

7.7 Native speaker results for Hypothesis 6

Hypothesis 6 proposed that the predicted effects of Hypotheses 1-5 would be greater for tasks that are more cognitively demanding, that is to say, that the effects on the decision and narrative tasks would be greater than the effects on the personal task.

By returning to the figures presented for the native speaker fluency measures in table 7.1 above we can see that the effect of planning was least (and always statistically insignificant) on the personal task, the one considered to be the least cognitively taxing. One obvious explanation for this is that the personal task offered the least scope for increased fluency because performance even in the unplanned condition was relatively fluent to begin with. On the narrative and decision tasks the difference in the mean total of pauses between the no planning and planning conditions is significant and considerable (20.25 to 5.67, $p = .000$ for the narrative and 28.25 to 14.80, $p = .000$ for the decision). The difference in the mean total of silence between the no planning and planning conditions is also significant and considerable (40.88 seconds to 7.17 seconds, $p = .001$ for the narrative, and 41.38 to 20.00, $p = .000$ for the decision.) The fact that the difference is proportionately greatest for the narrative could suggest that this task was found to be the most difficult by the native speakers and thus benefitted most from the extra attention afforded by planning time. Overall the mean scores for pausing and total of silence support Hypothesis 6.

The native speaker results for the other fluency measures of false starts, replacements, repetitions and reformulations do not achieve significance except for false starts in the narrative and decision tasks, and replacements only in the narrative. It is therefore not possible to reach any robust conclusions about support for Hypothesis 6. However, it is worth noting that across all but one of these fluency measures the planning condition has least impact on the personal task and more impact on the narrative and decision tasks, suggesting a trend in favour of support for Hypothesis 6.

For complexity, there is strong support for Hypothesis 6. The difference between the no planning and planning conditions was smallest and statistically insignificant for the personal task (1.20 vs 1.30, $p = .09$ *ns*). For the narrative and decision tasks the difference was greater and highly significant (1.24 vs 1.62 $p = .002$, and 1.36 vs 1.76 $p = .000$). The very small difference in the no planning and planning conditions on the personal task, a mere 0.1 of a clause, can be seen as a reflection of the limited linguistic demands of the task (giving directions) as well as its limited cognitive demands (dealing with very well known information). The appreciably larger difference between the no planning and planning condition for the narrative and decision tasks, 0.38 and 0.40 clauses respectively, can be seen to reflect the comparatively greater linguistic demands of these tasks (telling a story and expressing opinions) as well as the greater cognitive effort demanded by dealing with unfamiliar and/or unpredictable information. The results do not suggest, however, contrary to the results obtained for the number of pauses and total of silence, that the subjects found the narrative task more taxing than the decision task.

Overall we can conclude that the native speaker results strongly support Hypothesis 6 on complexity and offered partial support on fluency.

7.8 Non-native results for Hypothesis 6

The fluency measures of pausing and total of silence shown in table 7.3 above offer some support for Hypothesis 6. The difference between the number of pauses in the no planning condition and the number of pauses in the two planning conditions is small for the personal task (19.2 vs 10.0 and 11.2) compared to the difference in the narrative (30.3 vs 15.3 and 8.0) and the decision tasks (37.0 vs 17.3 and 17.3). Similarly for the total of silence, the mean total for the personal task reduces from 31.8 seconds in the no planning condition to 19.5 and 14.5 seconds in the two planning conditions, while the reduction for the other two tasks is considerably greater (120.3 vs 29.3 and 14.2 seconds for the narrative task, 91.4 vs 25.8 and 29.5 seconds for the decision). In the case of the narrative task, these differences are particularly impressive. This, coupled with the fact that the detailed planning condition had its greatest impact on this task, is evidence that the non-

native subjects found it to be the most taxing. Except for repetitions and replacements in the decision task, the other measures of dysfluency did not achieve significant F values in the one-way ANOVAS, and the mean scores, presented in Table 7.7 below reveal no consistent trend. Accordingly, Hypothesis 6 does not receive any further support

Table 7.5 showed the results for complexity, as measured by the number of clauses per AS-unit. The one-way ANOVAS showed that planning produced significantly more complex language on all three tasks. They also reveal that it was the personal task where this effect was smallest, with a mean 1.11 clauses per AS unit in the no planning condition rising only slightly to a mean of 1.16 and 1.26 in the two planning conditions. Although detailed planning had a stronger effect than undetailed planning, the overall increase in complexity compared to the no planning condition was only 0.15 of a clause. In the case of the other two tasks, the detailed planning condition produced greater complexity than the undetailed planning condition, (1.20 vs 1.43 and 1.68 for the narrative, 1.23 vs 1.35 and 1.52 for the decision) but the overall increase compared to the no planning condition was larger than that seen for the personal task: 0.48 of a clause for the narrative and 0.29 of a clause for the decision. Accordingly Hypothesis 6 is upheld for the non-native speaker results for complexity: planning time had a greater impact on the more cognitively demanding tasks. In line with the results for fluency, it would appear that the narrative task was found to be more taxing than the decision task.

A different picture emerges for the results for accuracy, as shown in Table 7.6 above. Here the narrative task, which the results for fluency and complexity have indicated was the most taxing, does not show the predicted increase in accuracy for the planning conditions. The F value from the one-way ANOVA did not reach significance in the narrative task, and the mean scores show that although there was an increase in accuracy between the no planning and detailed planning condition for each of the three tasks, this increase was smallest in the narrative and greatest in the personal task (61% vs 65% and 64% vs 76%). Moreover, although the impact of detailed planning is the greatest for the narrative task, it is in the opposite direction, i.e. less accuracy, not more. It is the

cognitively least taxing personal task therefore that shows the greatest effect for planning and accordingly Hypothesis 6 is not upheld for accuracy measures in the non-native speaker data.

For Hypothesis 6 therefore we can say that the non-native speaker results offer strong support for the complexity measure of subordination, partial support for the fluency measures and no support for the accuracy measures.

7.9 Results for Hypothesis 7

This was the null hypothesis, proposing that there would be no difference in the way planning time affected native and non-native speakers in terms of the fluency and complexity of the language they produced during the tasks. Lexical choices, the other aspect of performance covered by this hypothesis, will be considered in chapter 8. In this section we will look first at fluency and then complexity.

If we compare the fluency scores from Tables 7.1 and 7.3, we can see that although the non-native speakers always paused more often, planning time resulted in fewer pauses in all three tasks for both native and non-native speakers. The difference was statistically significant for all cases apart from the native speakers in the personal task. If we take the decision task as an example, we see that the non-native speakers paused an average of 37.0 times in the no planning condition and 17.3 times in both of the two planning conditions. The native speakers paused an average number of 28.25 times in the no planning condition and only 14.8 times in the planning condition. Similarly, for the total of silence, planning time reduced the mean totals of both native and non-native speakers across all three tasks, with only the personal task for the native speakers failing to reach significance. If we take the narrative task as an example, we see that the non-native speakers produced an average of 120.3 seconds silence in the no planning condition, and only 29.3 and 14.2 in the two planning conditions. The native speakers produced 40.88 seconds silence in the no planning condition, and 7.17 seconds in the planning condition.

The totals for silence are all much longer in the non-native speaker condition compared to those in the native speaker condition, and the scope for reduction in the planning condition is therefore greater. Nevertheless, for number and length of pauses, the null hypothesis is upheld.

As we have seen in section 7.2 above, in the non-native speaker data there was an unexpected and significant increase in the number of repetitions and replacements under planning conditions in the decision task (9.9 vs 20.1 and 17.9 repetitions $p<.05$, .056 vs 1.38 and 2.63 replacements $p<.05$). The ANOVAs showed no significant effect for planning on false starts and reformulations in the decision task, but the mean scores show that planning maintained or increased the incidence of these measures, and did not reduce them. (9.50 vs 13.63 and 10.25 false starts, 3.50 vs 4.88 and 3.87 reformulations). Equally, though the ANOVAs reveal no significant effect for planning on these measures of dysfluency for either the personal or the narrative task, the mean scores show these were most often higher in the planning conditions. Table 7.7 below makes this clear.

Table 7.7: mean scores for non-native measures of dysfluency

	repetitions		
	personal task	narrative task	decision task
no planning	12.06	7.19	9.94
undetail planning	11.38	10.63	20.13
detailed planning	15.25	6.00	17.88
	ns	*ns*	$p < .05$

	reformulations		
	personal task	narrative task	decision task
no planning	2.00	2.06	3.50
undetail. planning	2.37	2.88	4.88
detailed planning	3.5	2.00	3.87
	ns	*ns*	*ns*

	replacements		
	personal task	narrative task	decision task
no planning	1.33	1.38	0.56
undetail. planning	3.25	1.88	1.38
detailed planning	1.88	2.40	2.68
	ns	*ns*	$p<.05$

	false starts		
	personal task	narrative task	decision task
no planning	8.00	7.06	9.50
undetail planning	7.00	7.13	13.63
detailed planning	7.00	5.40	10.25
	ns	*ns*	*ns*

If we compare this trend of results with those obtained for the native speakers, detailed in table 7.1 above, we see a difference. For the native speakers, the planning condition brought about a *reduction* in the number of repetitions and replacements across all tasks, (not significant for repetitions, significant only for replacements on the narrative). There was also a reduction in the number of false starts in the native speaker data under planning conditions, reaching statistical significance in the narrative (5.63 vs 2.59 $p = .04$) and the decision tasks (8.0 vs 4.07 $p = .02$). Reformulations were also reduced by planning in both the narrative and decision tasks. It is only with reformulations in the personal task that these measures of dysfluency increase in the planning condition (very insignificantly from 1.0 to 1.2, $p = .65$). Planning time is therefore affecting the language of the native and non-native speakers in different ways, causing the non-native speakers to repeat and replace words more often while enabling the native speakers to speak more fluently. For replacements and repetitions in the decision task, the null hypothesis is not upheld. For the other tasks and other measures, there is insufficient evidence to reject the null hypothesis, but definite suggestions of a trend away from it.

For the measure of complexity, the effect of planning time on native and non-native speakers was similar. In neither was there any effect on the incidence of passives. For both, planning time increased the levels of subordinated language across all three tasks,

failing to reach significance in only the native speaker personal task. For the native speakers this effect of planning time was increased subordination in the decision task from a mean of 1.36 to 1.76 clauses per AS unit ($p < .000$) The effect was almost the same in the narrative where the mean increased from 1.24 to 1.62 clauses per AS unit ($p = .002$). In the non-native speaker data the strongest effects are also seen in the narrative and decision tasks though it was only the detailed planning condition that produced increases in subordination of a similar order to that seen in the native speaker data (1.20 vs 1.43 vs 1.68 clauses per AS unit in the narrative, $p < .001$; 1.23. vs 1.35 vs 1.52 clauses per AS unit in the decision task, $p < .001$). Accordingly, though the non-native speakers seem to need the extra attention afforded by detailed planning time in order to reach levels of syntactic complexity in language similar to native speakers ,we can conclude that the null hypothesis is upheld.

For Hypothesis 7 therefore, we can say that the null hypothesis is upheld for syntactic complexity, and for number and length of pauses, but is not supported for the measures of repetition and replacements in the decision task. The other dysfluency scores do not reach statistical significance for the non-native speaker data but do suggest that planning time might be increasing the incidence of these features, in contrast to the native speaker data where the scores indicate that planning time reduces them.

7.10 Discussion of the results

A central point of this thesis has been that attention is not an unlimited resource, and if several things demand attention at the same time then choices have to made where best to channel it. This study was designed to shed some light on the way native and non-native speakers make these choices, and to measure the extent to which their performance on tasks of varying difficulty is changed by the extra attention afforded by planning time.

Considering first the native speakers' performance and task type, the results clearly show that the personal task produced the most fluent performance i.e. the lowest incidence of *all* the markers of dysfluency in both planning conditions. This is consistent with the idea

that speaking about things which are well known and require little or no analysis (such as describing the route to your home) imposes no particular cognitive burden and allows speech to proceed relatively smoothly. The planning condition on the personal task did result in slightly lower mean scores for each of the measures of dysfluency, indicating that there was some scope for the extra attention to decrease the need for repetitions, false starts etc, but this was never significant.

As evidenced in the low mean scores for syntactic complexity, (1.20 clauses per AS-unit in the no planning condition, rising to only 1.30 clauses per AS unit in the planned condition) the task of describing a route does not require complex syntax. Indeed, it could be said that to couch directions in complex subordinated language is not helpful to the listener who has to remember and follow the route. The almost complete absence of the passive voice from the personal task data is probably also due to its inappropriateness: directions along a route are more normal in the active voice. The relative simplicity of the language used by the native speakers in the personal task under both planned and unplanned conditions can therefore be seen as a function of the perceived task purpose, (to give clear and memorable directions) and is a further explanation for the fluency with which both planners and non-planners were able to perform it.

The results for the fluency and complexity measures indicate that the other two tasks were, as predicted, more attention-demanding. All the scores for the measures of dysfluency were higher for the narrative and decision tasks than for the personal task, and planning time was always associated with lower scores for these measures. In several places these differences reached significance. In the narrative task and decision tasks there was a significant reduction in the number of false starts (5.63 to 2.59 and 8.0 to 4.07) indicating that planning time had made the subjects more sure about what it was they wanted to say. The biggest difference was however in the number and length of pauses. For both these measures planning time was associated with very significantly lower scores (e.g a mean of 20.25 pauses in the unplanned narrative compared to a mean of only 5.67 pauses in the planned narrative, $p = .000$) This suggests that without the extra attentional

capacity afforded by planning time, the native speakers coped with the demands of the narrative and decision tasks most often by frequent pausing, and to a much smaller extent by repeating or replacing words, and by reformulating or abandoning utterances.

At the same time as pausing often and at length during their narrative and decision tasks, the native speakers in the unplanned condition chose relatively simple, unsubordinated syntax with which to express themselves (a mean of 1.24 clauses per AS unit in the narrative and a mean of 1.36 in the decision). With the benefit of planning time, the syntax was very significantly more complex (1.62 clauses in the narrative, and 1.76 in the decision). This can be interpreted as reflecting both the greater opportunity for analysis and organisation of ideas available to the planners, and also the greater scope for complex expression afforded by story-telling and opinion-giving. The greater scope for complexity of expression did not, however, extend to a significantly higher incidence of the passive voice in these two tasks. Although there were more passives in the narrative planned condition, there were fewer in the decision planned condition. This suggests that the absence of the passive voice is not a good indicator of a lack of attentional resources, but (as noted above with regard to the personal task) probably has more to do with appropriacy of style.

There is some evidence in the native speaker data to indicate that the speakers found the decision task more taxing than the narrative. The scores for the measures of dysfluency (with the one exception of replacements) are higher for the decision task than for the narrative task in the no planning condition. Planning time is associated with an overall reduction in dysfluency measures, but these are still higher for the decision task than for the narrative. In the case of pausing and total of silence, the difference between the two tasks in the planning condition is comparatively wide (a mean of 5.67 vs 14.80 pauses and a mean total silence of 7.17 vs 20.0 seconds) suggesting that the native speakers were able to benefit more from planning the narrative than from planning the decision task, perhaps because the latter was more cognitively taxing. This is not necessarily due to the subject matter. It must be remembered that the narrative was essentially a monologic task

in which speakers were given an extended and uninterrupted turn. The decision task, by contrast, was discursive and required the participants to argue a point of view. For the native speakers, then, taking account of a partner's opinions is a relatively unpredictable matter and therefore can be seen as more taxing and less amenable to planning.

These results show that language choices are affected by the type of task a speaker is given to do. Relaying well-known information to someone who needs to remember and follow it is done in relatively unsubordinated syntax, telling a story or arguing a case is done in relatively subordinated syntax. These differences remain even when planning time is available. Contrary to the conclusions drawn by Ochs (1979) and Givon (1979) planned language is not always more complex than unplanned language, and certainly the low incidence of the passive voice throughout the data cannot be taken as an indication of insufficient attentional capacity.[27] Task type also impacts on the fluency with which it is transacted. The personal task, which involved well-known and well-structured information, was carried out with relative fluency in both planned and unplanned conditions. The narrative and decision tasks, by contrast, were performed with considerably more fluency when planning time was available. Without this extra attention, these tasks caused the native speakers to pause significantly more often and at greater length, as well as to abandon significantly more unfinished utterances.

The results have also shown that, whatever the task, planning time did make language performance more fluent and more complex. This was to a very small and insignificant degree in the case of the personal task, but to a greater and highly significant degree in the other two tasks. This effect for planning time, whether small or great, supports an information processing model of language production in which achieving fluent and appropriately complex language depends on the amount of attention available. It shows that, when attention is limited and the task demands are high, speakers will use relatively unsubordinated syntax and relatively frequent pausing and self-repair. These observations

---

[27] Indeed, the choice of, for example, 'It's been broken' over 'I've broken it' is clearly a matter of the speaker's intention, not the speaker's attention.

are important to bear in mind when considering the non-native speaker results because, in addition to the attentional demands imposed by the different task types, these speakers have the extra burden of performing in a second language over which they have more procedural than automatic control.

7.10.1 Task difficulty and task performance

Regarding task difficulty, it is clear that the non-native speakers also found the personal task to be the least taxing. As in the native speaker data, the scores in the no planning condition of the personal task show by far the smallest number of pauses and the smallest total of silence. (For example, there was a mean total of 31.8 seconds silence in the no planning condition of the personal task as compared to the enormously larger totals of 120.3 seconds in the narrative and 91.4 seconds in the decision task.) As in the native speaker data, the personal task without planning time resulted in the least subordinated language (a mean 1.11 clauses per AS unit vs 1.20 and 1.23) and, perhaps unsurprisingly, the most accurate (64% vs 61% vs 63%). But it is worth noting that the cross task range in these scores is very much narrower than the cross-task range in the fluency scores (only 0.08 for complexity and a tiny 0.03 for accuracy) indicating perhaps that the non-native speakers in the no planning condition were being deliberately unambitious in the narrative and decision tasks because they needed more than their frequent long pauses to produce language that was adequately complex and accurate, i.e. they needed planning time. Indeed, the impact of planning time on complexity is much greater in the decision and narrative tasks than it is in the personal task, (the cross-task range in detailed planning complexity scores is 0.42). This is consistent with the native speaker results and supports the suggestion that speaking about familiar information with a well defined structure, which is essentially the nature of the personal task, is something that can be adequately accomplished without the need for the extra attention to language afforded by planning time.

For the other two tasks, the no planning condition resulted in language that was, compared to the language of the personal task, very considerably less fluent, only slightly more complex, and slightly less accurate, all indications of a greater cognitive load. In contrast to the native speakers, however, the non-native speakers appear to have found the narrative task most taxing. Without planning, they paused at greater length and spoke with less complexity and more inaccuracy on this task than on the decision task.

It is interesting to speculate why this should have been the case. A prerequisite for accomplishing this task was a reasonably good imagination that could fit a series of pictures into a storyline, and there is no reason to suppose that the native speakers were better endowed in this regard than their non-native counterparts. One explanation might be that the non-native speakers were searching for a storyline that could be expressed within the limits of their vocabulary. (It would have been no use to have come up with an imaginative story and then been unable to explain it). This extra burden could have slowed them down and caused them to produce relatively simple and relatively inaccurate language. Additionally, the non-native speakers' poorest performance on the narrative could have been due to the relative unpopularity of the task as evidenced by their comments in the transcripts. Many of the non-planners seemed unsure of how to proceed or complained that they did not know what to do. Those with planning time, by contrast, did not voice such complaints and were all able to some degree to come up with a storyline for the pictures in language that was more complex and fluent than the language they produced in the planned decision task. The impact of planning time, therefore, was greater for the narrative than for the decision.

7.10.2 The overall effect for planning

If we look at the overall effect of planning time on fluency and complexity in the non-native speaker data, it is clear that this is generally parallel to the effect observed in the native speaker data. Complexity, as measured through the level of subordination, increased with planning time across all three tasks. (This is consistent with the results obtained by Crookes (1989) who showed that for non-native speakers undetailed planning

time is associated with more subordinated language). Moreover, it appears that those subjects with detailed planning were helped by the suggestions given to them in the planning stage to produce language that was consistently the most complex. Bearing in mind that the narrative task was apparently regarded as the most difficult, it is interesting to see that the detailed planning condition here resulted in the highest score of all for subordination (an impressive mean of 1.68 clauses per AS unit). It could be that the detailed planning condition not only allowed subjects to think up ideas to connect the pictures (which the non-planners apparently found very hard to do), it also allowed them to produce language that was suitably complex to express those ideas.

Fluency, as measured through pausing and total of silence, also significantly increased with planning across all tasks, though in contrast to the results for complexity, the detailed planning condition did not always result in the most fluent performance. The predicted monotonic relationship between planning condition and fluency only held up in the narrative task, where detailed planning enabled not only the most complex performance (as we have noted above) but also that with the fewest pauses and smallest total of silence. An almost monotonic relationship between planning and fluency is seen in the personal task, though the detailed planners produced very slightly more pauses than the undetailed planners (10.00 vs 11.5). In the decision task, the detailed planning condition does not reduce pauses or silence further than the undetailed planning condition, it maintains the former and increases the latter (17.25 vs 17.25 and 25.75 vs 29. 50). We may explain this by reference to the nature of the tasks. The personal and narrative tasks are essentially monologues, while the decision task is a discussion. Discussions need to take account of the contributions of the interlocutor, and although the detailed planning condition suggestions advised the subjects to try to anticipate what their partner might say, this was obviously not something which could be easily planned for. As noted above, a similar pattern was seen in the native speaker data, where the planning condition had less of a fluency effect in the decision task than in the narrative and personal tasks, and for which the same explanation was offered.

Though planning reduced the incidence of pausing and amount of silence in the performance of both native and non-native speakers, there was not a parallel effect on the other measures of dysfluency. For the native speakers, planning time was associated with a smaller number of repetitions, reformulations, replacements, and false starts in all three tasks (apart from the personal task reformulations where there was a very small and insignificant increase). For the non-native speakers by contrast planning time is generally associated with similar or higher scores for these measures, reaching a statistically significant increase in repetitions and replacements in the decision task. To help explain this, it is useful to distinguish between pausing and repetition as an indicator of forward planning (i.e. thinking about what to say next), and reformulations, replacements and false starts as indicators of retrospective repair (i.e. thinking about what has just been said). Planning time seems to decrease the need for the non-native speakers to pause in order think about what to say next, and instead allows them to use the more natural time-gaining device of repeating what they have just said. Planning also seems to increase the likelihood that non-native speakers will reflect on what they have just said and see ways of improving it through reformulation or replacement, or else decide that the best course is a false-start and a new utterance. This would indicate that, in the no planning condition, where attentional resources are especially limited, the non-native speakers are more likely to be silently thinking about what to say next, than they are to be reflecting on what they have just said. It must be remembered, of course, that such an analysis is not based firmly on statistical evidence but rather on an observable trend in the results. Clearly these matters deserve further investigation.

7.10.3 The effect of planning on accuracy

So far we have been considering how the non-native speaker results are similar to or different from the native speaker results. It is important now to consider a dimension of language performance that was not calculated for the native speakers, i.e. accuracy. As noted above in 6.1 it was expected that the accuracy of the native speaker's language would not be affected by the planning condition, and indeed it was not. A reading of the transcripts revealed some non-standard grammar and some wrongly selected lexical items

in both planned and unplanned conditions (discussed further in section 9.1) but the very low total of nineteen. such errors, more or less evenly distributed across the planned and unplanned conditions, meant that a statistical analysis would not have been possible. For the native speakers clearly, grammatical accuracy was not something affected by attentional capacity. For the non-native speakers, however, planning had a marked effect upon accuracy, though not in the way predicted by hypothesis 5. It was proposed that the extra attention afforded by undetailed planning would result in more accurate language than that found in the no planning condition, and that the even greater attention supposedly provided by detailed planning would result in the most accurate language of all. In fact, the results show that it was consistently the undetailed planners who produced the most accurate language in all tasks, and in the case of the narrative task the detailed planners produced the least accurate language of any group of any task.

In order to account for this unexpected pattern of results, it is helpful first to recall that previous research into the effect of planning on accuracy is not consistent. Ellis (1987) did find that planning time increased accuracy, but this was by measuring only the past tense and unfortunately confounded speech with writing. Crookes (1989) who investigated speech only, found that (undetailed) planning time did not increase accuracy, although he looked only at specific linguistic items and may therefore have failed to detect a global effect. The present study, by calculating error-free clauses, used a measure which was sensitive to a more general improvement and which, in fact, revealed such in the undetailed planning condition. We may say therefore, that as far as undetailed planning is concerned, accuracy does increase with greater attentional resources, just as fluency (mostly) and complexity also increase with greater attentional resources. The question we must address now is, why does this improvement continue in complexity and fluency (mostly) but not in accuracy when planning is aided by detailed suggestions.

What seems to be happening is that in the undetailed planning condition improvements in accuracy and complexity (as well as in fluency) can be achieved within the limitations of attentional capacity, but these performance features enter into competition with each other

in the detailed planning condition. In other words, the detailed planning condition lead to a level of performance pressures that were too great for the attentional resources available, and the speakers therefore had to choose which aspect of performance to prioritise. Complexity emerged as the pre-eminent goal, at the expense of accuracy, and to some extent of fluency also. In the personal and decision tasks detailed planning caused complexity to peak, while accuracy reduced to a level somewhere between the no planning and undetailed planning conditions. In the narrative task, detailed planning took complexity to the highest level of all, while accuracy dropped to below the level of the no planning condition. There was thus a clear tradeoff between complexity and accuracy, with fluency dividing somewhat between the monologic personal and narrative tasks (where it increases in line with complexity) and the discursive decision task (where it appears to decrease a little in line with accuracy). In order to explain the excessive pressures on processing resources that the detailed planning condition seemed to produce, one might say that the detailed planning condition, in which suggestions were given on how to prepare for the task, lead subjects to be over-ambitious. They may have used planning time to formulate complex ideas that they wished to express, which in turn required syntax and vocabulary beyond their level of comfortable control, or else which required so much in the way of attention to execute that maintaining accuracy was not possible. We cannot conclude from these results precisely how the non-native speakers were using their detailed planning time, nor precisely how their attention was allocated during the subsequent tasks. It is only possible to conclude that they were not able to show simultaneous and equal improvements in complexity, accuracy and fluency.

7.11 Summing up

Though the patterns of results are in places rather complex, we can nevertheless broadly conclude that the overarching hypothesis that language is affected by attentional demands and attentional resources has been upheld. The different cognitive loads of the three tasks showed clear effects on the performance of both the native and non-native speakers, and the easing of that load through planning time was demonstrated in both by increases in

fluency and complexity. The burden of being accurate and complex in L2 performance was also shown, with clear indications that a) these two performance features enter into competition with each other for attentional resources and b) when this happens, complexity is prioritised. Fluency has been shown to be a rather complex feature, dividing for the non-native speakers into the forward-planning pauses (which reduce significantly with planning time) and the backward-reflecting reformulations, replacements and false starts (which do not). For the native speakers, all measures of dysfluency decrease as the attentional load is reduced through planning time. What remains now is to investigate the lexical dimension of planned and unplanned task performance, to see to what extent this too reflects the amount of attention available to a speaker. This is the subject of chapter eight.

Chapter Eight

Lexical results and analysis

Hypothesis 3 predicted that if planning time is given before a task, a speaker will rely less on pre-formed lexicalised language to express ideas. It was reasoned that the use of such lexicalised language is a time-gaining processing strategy that will be most evident when attentional resources are limited. This chapter therefore will discuss the argument that much of language use can be accounted for by memorised 'chunks' rather than the operation of syntactic rules, and will then discuss how such chunks can be defined, and identified in the data. The results of the lexical analysis of the data are then presented, followed by a discussion on how these lexical results compare to the results obtained from the syntactic and fluency analyses.

8.1 The role of memory in speech

There is now a body of research in linguistics which studies the extent to which words operate in fully or partially fixed combinations as opposed to within a productive system of syntactic rules (see Weinert 1995 for an excellent review, also Howarth 1998). The use of such memorised sequences is described as a processing strategy for first and second language use which permits fluent and fast language production (Raupach 1984, Pawley and Syder 1983). As discussed in chapter two, building a store of memorised sequences has also been suggested as a learning strategy adopted by both first and second language learners whereby regularly encountered combinations of words are committed unanalysed to memory and then analysed for productive grammatical regularities (Ellis 1996).

This contrasts sharply with the classic Chomskyan model of linguistic competence (Chomsky 1957, 1965). In this, the complexity of language and the speed with which

young children nevertheless master it can only be explained by positing an innate language faculty, endowed with the principles of a Universal Grammar. These principles, 'pre-wired' in the brain, enable the otherwise cognitively helpless infant to work out the syntactic structures of the language(s) it is born into. Native speakers thereby acquire an impressive and faultless knowledge, most of it below the level of consciousness, of the set of rules that governs their language. It is this syntactic knowledge, together with a lexicon of tens of thousands of words, which forms the basis of language proficiency. It enables speakers to recognise which combinations of words are grammatical and which are not, and gives them the potential to generate a infinitely large number of well-formed utterances.

Chomsky's model has been widely accepted, though it remains controversial. It cannot, for example, account for why native speakers do not fully exploit their linguistic competence. Becker (1975) was one of the first to observe this, pointing out that, although native speakers might possess grammatical and lexical knowledge that enables them to produce a potentially infinite number of novel well-formed utterances, in practise they do nothing of the kind, preferring to cobble together memorised groups of words that everyone has heard before. He called this bank of memorised sequences a 'phrasal lexicon'.

> We start with the information that we wish to convey and the attitude towards that information that we wish to express or evoke, and we haul out of our phrasal lexicon some patterns that can provide the major elements of this expression.....Then the problem is to stitch these phrases together into something roughly grammatical, to fill in the blanks with the particulars of the case at hand, to modify the phrases if need be, and if all else fails to generate phrases from scratch to smooth over the transitions or fill in any remaining conceptual holes. (p.28)

Essentially the same observation has been made many times since. (See, among others, Pawley and Syder 1983, Widdowson 1989, Willis 1990, Nattinger and Decarrico 1992, Sinclair 1991, Lewis 1993). The fluency and familiarity of native-like language is explained by the fact that it is generally *not* composed of novel combinations of words but instead uses a lot of prefabricated sequences shared by everyone in the speech community.

The development of computer software that can carry out lexicographical analyses of large databases of language has enabled researchers to collect hard evidence of the remarkable pervasiveness of fully and partially preconstructed elements in both the spoken and written language. The Cobuild project which was one of the first extensive lexicographical analyses of the English language, led its director Sinclair (1991) to conclude that language use was guided in the main by 'an idiom principle', and was best described as streams of collocational patterns that flow into each other. On the evidence of computer analysis we are not really creative when we compose language, preferring to use words in the same sequences and combinations as everybody else.

This does not mean that we are doomed to speak in clichés (though some people do more than others). Clearly we all have knowledge of grammar and use it to fit together familiar sequences of words, as well as to construct entirely novel ones. From the perspective of psychology, Ellis's (1996) model of language learning, based on our often under-estimated capacity to detect and remember patterns, elegantly accounts for both our creative knowledge of language rules and our store of ready-to-use word sequences. Ellis argues that both first and second language learning is essentially the acquisition and analysis of memorised sequences. A learner acquires sequences of sounds in order to learn words and sequences of words in order to learn phrases. In both cases acquisition follows repeated exposure to examples. The more often certain sounds are heard in the same sequence the more likely that sequence is to be transferred to long-term memory. The more often words are encountered in particular patterns, the more likely those patterns are to be stored in

long-term memory.[28] The learner acquires a database of words and their relationship to the other words that regularly occur with them. In the case of the first language learner, there is an automatic and implicit analysis of these relationships that results in the abstraction of grammatical regularities, but this does not override the memory store. We don't select the rule-based but peculiar *this road is main* over the routine and familiar *this is a main road.* In this view, it is not necessary to explain first language acquisition by reference to an innate set of linguistic principles. Contrary to Chomsky's model of pre-wired, innate linguistic principles, grammatical knowledge is the *outcome* of language use, and not the other way round.

Consequently, we are able to process language in two ways: the syntactic rules abstracted from memorised exemplars are used for the comprehension and/or production of novel or complex structures, while for the comprehension and/or production of routine and familiar structures we use more our instant access to a memory store of fixed or partially formed phrases. The time available and our familiarity with the subject matter are important determinants. The less time we have to prepare what we are going to say and decode what we have just heard (and in normally fluent conversations we have little) the more we exploit memory. The more familiar we are with the subject matter, the more likely it is that our memory will contain relevant ready-to-use language and the faster we are able to process it. The opposite case is equally true. Anyone who has tried to follow complex instructions for an unfamiliar task (such as tuning a remote control to a new television), will know that a slow and repeated word-by-word analysis of the language is often necessary, even if none of the words themselves are unknown. Ellis (1996:116) puts it succinctly: 'linguistic analysis is as deep as necessary and as shallow as possible.'

This observation accords well with the model of the brain as a limited-capacity information processor, discussed at length in chapter two (e.g. Anderson 1982, McLaughlin 1987) As we saw, paying attention to both the form and meaning of language represents a

---

[28] There is good evidence that the ability to remember and repeat sequences of sounds (phonological short-term memory) is an accurate indicator of language learning aptitude (Bley-Vroman and Chaudron 1994).

considerable cognitive burden. If we had to process all words by reference to the rules of grammar, we might have insufficient capacity to attend to the content of their message, especially when the forms are complex and the message unfamiliar. We might often have to pause for thought, or even to stop everything else we are doing, in order to carry on a conversation. Our reliance on ready-to-use 'chunks' of one kind or another is a useful processing strategy, enabling normal speech to be processed more or less seamlessly.

8.2 Defining memorised sequences

Although there seems to be a general consensus as to the reality of these prefabricated sequences, they are difficult to define concisely. They take many forms and have many names (Weinert 1995, Becker 1976). There are fixed phrases such as *'be that as it may' 'see you later' 'when all is said and done'*, sayings such as *'there's no time like the present' 'dead men don't tell lies'*, metaphors such as *'ignorant as dirt, proud as a peacock,'* collocations such as *'flatly refuse' 'catch a cold' 'fail miserably'*, phrasal verbs such as *'take after' 'get along with'*, *'look forward to'* and, perhaps least obvious but nevertheless widespread, partially preconstructed sentence stems such as *'to bring something to the attention of someone' 'to be sorry to keep someone waiting'*. The common feature shared by all is that they comprise or contain elements produced as wholes. Pawley and Syder note that conversational speech is characterised by fluent, grammatical clauses of four to ten words delivered at a faster than normal rate of articulation, and claim this as further evidence that the stream of speech is constructed of fully or partially memorised chunks that are single choices, not individual words brought together for the occasion.[29]

---

[29] A Chomskyan might argue that the use of such chunks is simply a *performance* feature, a necessary strategy for the realisation of *competence* under the time-pressure of speaking. But these chunks of language are also evidenced in written data, where there is no such excuse. Moreover, the fact that all members of a speech community appear to share the same language chunks means that they are not one-off solutions to particular performance demands but are themselves an important part of any native speaker's knowledge.

Whereas many lexical sequences are produced as invariable chunks, (such as *in the end, never mind, what's the matter*) and many others are written as single words (such as *painstaking, nevertheless, always*), others are open to some degree of analysis and substitution of elements. Consider, for example, the following sequences which appear to act as discourse organisers by indicating the general nature of what is about to be expressed.

I was (just) wondering if.....

Yes, but the point/fact is.....

you'll never guess/believe.....

what really gets/bugs/kills me is.....

that is/may be all right/OK/fine for some people but...

Perhaps less obvious, but still powerful, are the bonds between words that work at greater distances than immediate proximity, as in the case of partially preconstructed phrases, sometimes called sentence stems. In these, a ready-made framework is provided into which the speakers have to fit only the particulars of the precise meaning to be expressed.

who the hell do you think you are?

who the devil does she think she is?

who the blazes do they think I am?

She's not the sort of person who would go around murdering people.

Am I the sort of person who would go around stealing milk bottles?

He's exactly the sort of person who would go around looking for trouble.

She's only gone and lost my keys!

You've only gone and broken it!

He's only gone and got himself arrested!

They've only gone and got married!

As we have noted, there are various terms (some overlapping) for the different ways words link up together: lexical phrases, phrasal lexemes, chunks, sentence stems, multi-word items, formulaic phrases, and many others. The analysis of the data in this research study is concerned with the whole range of these, and for simplicity's sake the term *'lexicalised'* is used to describe any combination of words which are stored in memory as a fully or partially formed sequence, as opposed to words that are brought together on a particular occasion. In this way, it is hoped to cast as wide a net as possible.

8.3 Coding the data.

This was somewhat problematic, and needs to be discussed in detail. The coding could not be accomplished by a computer analysis for the very good reason that although computers can work with huge databases of language quickly, they cannot distinguish language which is lexicalised from language which is not, except for any fixed phrases and collocations which occur repeatedly. Unfortunately, the majority do not occur repeatedly even in huge corpora. Moon (1998) reporting on an analysis of a 118 million word database of English finds that whereas phrasal lexemes (her word for the whole range of fixed and semi-fixed word sequences) were very common, most individual examples occurred less than once every million words. Even a corpus as large as The Bank of English at the University of Birmingham, now nearly three hundred million words, fails to show even a single example of many phrases that would be considered a normal part of any native speaker's repertoire. This indicates that the memory store of lexicalised sequences is a) vast, b) retentive, and c) made up in great part by items which are infrequently called upon. It also means that in a relatively small corpus most lexicalised sequences are very unlikely to occur more than once, and therefore a computer analysis which uses frequency as a criterion for identifying lexicalised language is not going to be helpful.

The approach adopted in the analysis of this data was to exploit native speaker intuition because, unlike a computer analysis, this can recognise even rarely used lexicalised

sequences.[30] However, native speakers, especially those who are not accustomed to a conscious analysis of language, can be inconsistent in their judgements. (Willis (in prep) reports uneven results). One way to avoid this problem is to use native speakers who are already well versed in applied linguistics, whose intuition is shaped by professional experience, and who therefore have a good understanding of what is required of them. This ought to give greater reliability of judgement.[31]

There is, of course, a further problem with analysing non-native speakers data for lexicalised language. Native speaker intuition cannot necessarily penetrate all the way into the phrasal lexicon of individual learners. Non-native speakers are likely to have memorised sequences which are peculiar to themselves (a kind of lexical idiolect) and thus unrecognisable to others unless flagged by frequent repetition. The problem of identifying these may be insuperable, but addressed to some degree by using native speakers with extensive English language teaching experience who are thus familiar with non-native varieties of English and can make an informed judgement as to whether and where lexicalised sequences occur.

Finally, as even applied linguists with extensive English language teaching experience can be unreliable, it was considered important to collect information from several, to collate it, and to accept as lexicalised only those word sequences which the majority have identified. In this way, dubious or borderline examples can be eliminated and rare but clear examples are not missed. It was felt that to have asked volunteer informants to analyse the transcripts of all three tasks (almost 60,000 words) would have been extremely onerous. Therefore, as a preliminary study into the use of lexicalised language by these four groups of speakers, only the data from the discussion task was analysed. The resulting corpus is thus only about 20,000 words (still a huge reading task), but has the unusual virtue of being recorded from speakers of approximately the same age and educational attainment,

---

[30] I don't think I have ever used the expression 'well, you could have knocked me down with a feather' and I cannot recall the last time I heard it or read it, but I have no trouble recognising it at once as part of the normal British English idiom.

all discussing the same topic under identical conditions. Most large corpora used for lexical analysis suffer somewhat from being made up of data from disparate sources. To some extent, what this corpus lacks in size it makes up for in uniformity.

8.3.1 Procedure for coding the data.

The seven native speaker informants who analysed the transcripts were all university teachers of Applied Linguistics with many years experience in English as a Foreign Language. All were given a complete set of the transcripts of the decision task, shuffled so that native speaker and non-native speaker, planned and unplanned conditions were randomly mixed. The informants were given the same instructions to read through the transcripts and, without consulting anyone else, to mark any language which they felt had not been constructed word by word, but had been produced as a fixed 'chunk', or as part of a sentence 'stem' to which some morphological adjustments or lexical additions had been required. The information provided by the seven native speakers was then collated onto a master document by bracketting each phrase every time a phrase was identified by an informant as lexicalised. The result was a set of transcripts marked like this:

> (((((((it doesn't matter))))))) (((((what the circumstances))))), (((((she didn't have the right to))))) (((((take his life))))). If she was that er emotionally ((((((you know))))))) er distressed, then she should have- ((((((I don't know))))))) (got out of the situation). (((((It's difficult to say))))) when you are not (((((in the situation))))) but ((((((at the end of the day))))))) she did (((((take another human life))))). (((((((There you go.)))))))

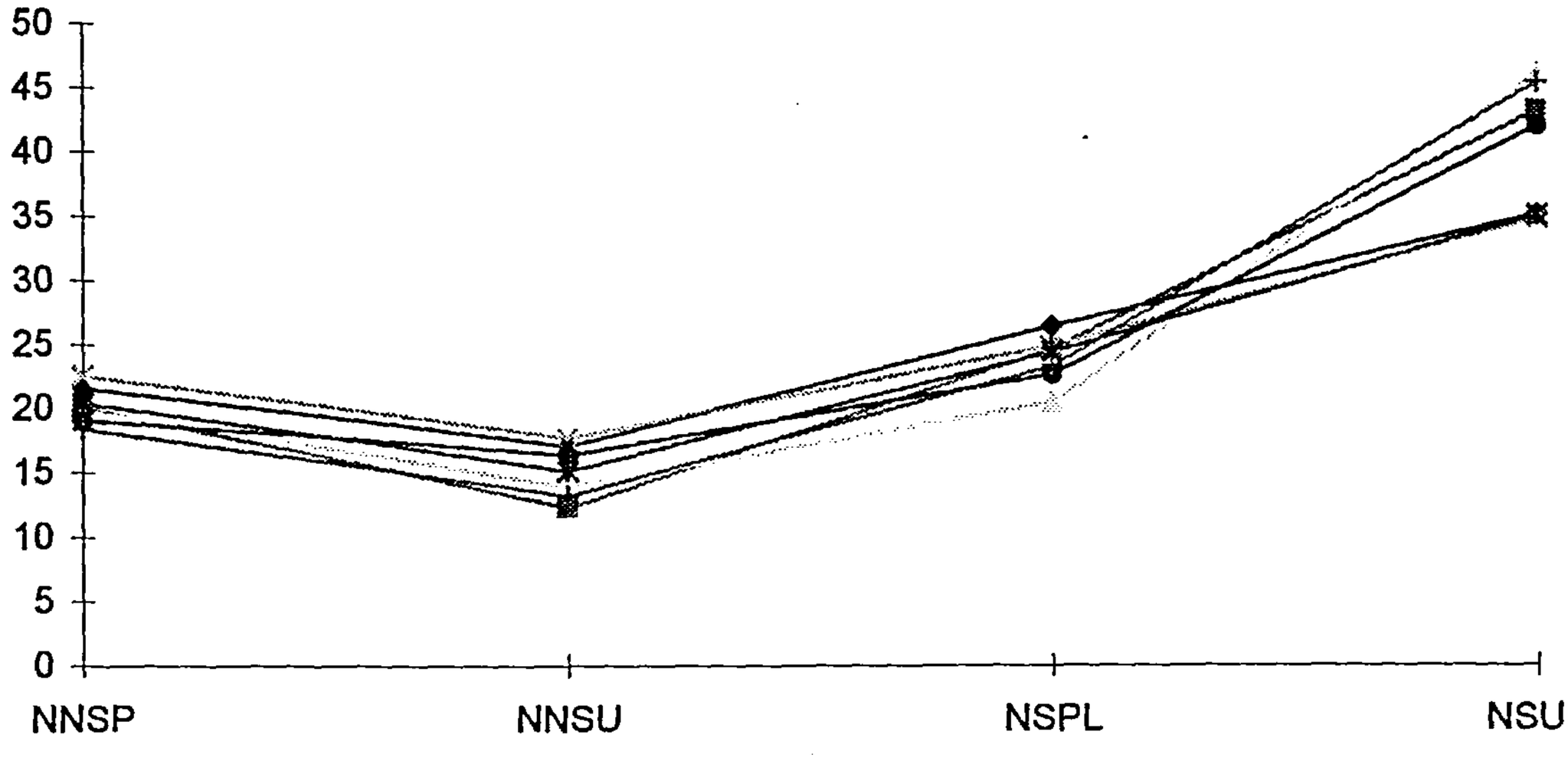
Only those phrases which were bracketted by five or more of the informants were used in the analysis

---

[31] . Using native speaker intuition has a long history in linguistic research. It is the preferred way of proceeding for Chomskyans.

According to the written comments of all seven informants, the coding was not an easy task. Lapses of concentration with reading meant missing even obvious examples of prefabricated language, so progress was slow and exhausting. All seven reported difficulty in knowing where exactly to mark boundaries of some lexical chunks and stems as one could overlap or even envelop another. Nevertheless, after a certain amount of self-imposed revision, each reported feeling reasonably confident with their coding and a comparison of their figures showed this confidence was not misplaced. Figure One below shows how the lexicalised language identified by each of the seven informants was distributed across the four conditions. As can be seen, all seven informants found the most lexicalised language in the native speaker unplanned condition, and all seven found the least in the non-native speaker unplanned condition. In fact, the graph shows considerable consistency in its curve at all points and indicates that the seven informants, working independently, were producing similar results.

Figure One : Distribution of lexicalised sequences across the four conditions
(expressed as % of each informant's total)



nnsp = nonnative speakers, planned; nnsu = nonnative speakers, unplanned;
nspl = native speakers, planned; nsup = native speakers, unplanned

8.3.3 Analysis of the data

For each of the four conditions (native and non-native, planned and unplanned) the number of lexicalised chunks which had been bracketted at least five times was counted. This meant that the results would report on only those words which five, six or all seven of the informants had intuitively judged to be lexicalised in some way. Also calculated for the four conditions was the number of words appearing inside brackets as a percentage of the total number of words produced by the speakers in that condition. Compound words (e.g. *breadknife*,) were counted as single items if so spelled in Webster's Dictionary. Others (e.g. *life sentence*) were counted as two. The analysis included both the fixed words in a lexicalised sequence as well as any required by the context to be added within it i.e. *set him free* and *set her free* both count as three words. This was not an entirely satisfactory way to proceed because informants sometimes disagreed as to where exactly the sequence boundaries were. (*It's a question of* vs *it's a question of asking*....).In these cases, the majority decision was accepted. There is a further difficulty, however. If lexicalised sequences are used and stored as wholes, it is not straightforward to justify counting the individual words into which they can be analysed. But in the absence of a better way to calculate the proportion of lexicalised language in the data, this preliminary analysis was undertaken, with reservations. Also calculated were the number of sequences which occurred only once, those which occurred between two and six times, and those which occurred more than six times.

8.4 Results.

Table 8.1: Amount of lexicalised language identified in the corpus

	total number of words produced	number of identified lexicalised sequences	number of words in inside lexicalised sequences	words inside lexicalised sequences as % of total words
NNS unplanned	4089	261	690	16.87%
NNS planned	5577	349	961	17.23%
NS unplanned	5584	621	1803	32.29%
NS planned	4544	342	1140	25.08%

NS = native speaker NNS = non-native speaker

The data is all from the first five minutes of the task. The figures represent the raw totals and percentages of words and lexicalised sequences for each of the four groups. It was not possible to collect totals for individual subjects, as was done for the syntactic and fluency measures. Therefore, given that only four figures were available for two speaker types and two planning conditions, no statistical procedures were attempted beyond comparing the totals and percentages for the four groups..

The results in table 8.1 show that, whatever the planning condition, the native speakers' language was composed of a much higher proportion of identified lexicalised sequences. The planning condition had a different effect on the language of the native and non-native speakers. With the advantage of planning, the non-native speakers produced a lot more language in these five minutes (5577 words compared to 4089 words), whereas the native speakers with planning produced less (4544 words compared to 5584 words). However, for the non-native speakers the percentage of these words occurring as part of lexicalised sequences is practically unchanged by the planning condition (16.87% compared to 17.23%), while for the native speakers the planning condition causes the proportion of lexicalised language to reduce considerably from about one third to one quarter. (32.29% to 25.08%). Taken together, these results suggest that native speakers are less verbose when they plan what they are going to say, and produce language which is less reliant on lexicalised sequences. For non-native speakers, the results suggest that planning enables them to speak more, but without affecting the level of identified lexicalised sequences in their language which is anyway rather low. Hypothesis 3 is therefore confirmed for native speakers, but not for the non-native speakers

A finer-grained analysis of the data noted those sequences which appeared often i.e. seven or more times[32], those which appeared between 2 and 6 times, and those which appeared only once. The results are shown in table 8.2 below.

---

[32] This is necessarily an arbitrary choice, as it is not possible to define what constitutes 'often' in this particular context. However, it was felt that it is reasonable to consider a sequence which occurs at least seven times in a four to five thousand word sample as appearing 'often'

Table 8.2: Frequency and variety of lexicalised sequences identified in the data

	number of lexical sequences identified	lexical sequences found only once			lexical sequences found 2-6 times			lexical sequences found 7+ times		
		no of types	no of tokens	*% of total*	no of types	no of tokens	*% of total*	no of types	no of tokens	*% of total*
NNS unplanned	261	65	65	*24.9*	29	85	*32.6*	4	111	*42.5*
NNS planned	349	56	56	*16.0*	37	100	*28.7*	10	193	*55.3*
NS unplanned	621	198	198	*31.9*	77	222	*35.7*	11	201	*32.4*
NS planned	342	190	190	*55.6*	33	81	*23.7*	4	71	*20.8*

This analysis reveals that the non-native speakers used some lexicalised sequences often. Of the 261 lexicalised sequences identified in the non-native, unplanned condition, a remarkable 42.5% is accounted for by just four recurring sequences.

I think (70 times),
I don't know (23 times)
how long (11 times)
all right (7 times)

With the planning condition, the non-native speakers show a similar dependence on a slightly larger number of lexicalised sequences. Fifty-five percent of their 349 identified lexicalised sequences were accounted for by ten frequently recurring phrases, the majority of which appear to function as fillers or discourse organisers.

I think (87 times)
you know (33 times)
I don't know (17 times)
I don't think (9 times)
I know (used to signal agreement, 9 times)
for me (used to begin an utterance, 9 times)
I mean (8 times)
what do you think (7 times)

a lot of (7 times)

real life sentence (7 times)

Interestingly, something similar is happening in the two native speaker conditions. In the no planning condition, 621 lexical sequences were identified of which 32.4% were accounted for by just eleven oft-recurring phrases. Some of these were the same as those relied upon by the non-native speakers, but most were not:

I think (51 times)

I mean (34 times)

I don't think (20 times)

you know (19 times)

sort of (16 times)

I don't know (15 times)

a life sentence (12 times)

I suppose (9 times)

set him/her free (9 times)

yeah, but (to begin an utterance, 9 tiimes)

or something (7 times)

. In the native speaker planning condition, only 342 lexical sequences were identified, and of these 20.8% were accounted for by four recurring sequences. Again the majority of these oft-used phrases are fillers and organisers:

I think (35 times)

you know (19 times)

I mean (9 times)

I don't know (8 times)

In the unplanned non-native condition, 24.9% of identified lexical sequences occurred only once. In the planned non-native condition this figure is only 16.0%., indicating much less variety in their selection of lexicalised sequences. There is an opposite trend in the native speaker conditions. Without planning, 31.9% of the identified lexicalised sequences occur only once. With planning this figure is much higher at 55.6%, suggesting a more varied selection of lexicalised sequences.

8.5 Discussion

We may sum up these results by saying that, whatever the planning condition, there was much more lexicalised language identified in the native speaker data than was identified in the non-native speaker data. Planning time not only reduced the proportion of lexicalised language in the native speaker data, it also reduced the repeated use of a small number of sequences while increasing the variety of the rest. For the non-native speakers on the other hand, though planning time increased the amount of language they produced it did not reduce their repeated use of a small number of sequences which accounted for around half of all the lexicalised language identified in their data. Moreover, and in contrast to the native speaker results, planning time seems to have produced a less varied range of lexicalised language in the non-native speaker data, possibly because non-native speakers have a smaller pool of lexicalised sequences to draw upon and the more language they produce the greater the chances are that sequences will be used again. However, this finding, as with all the non-native speaker findings, could be because the seven informants who coded the data were not able to recognise language that was un-native-like, but nevertheless lexicalised. This is a point we must return to below.

It is interesting to compare these results with the findings of the syntactic analysis of the data in chapter 7. The relevant scores for the discussion task are reproduced in Table 8.3 below.

Table 8.3: Mean accuracy, complexity and fluency scores for the Discussion Task

	non-native speakers			native speakers	
	unplanned	unguided planning	guided planning	unplanned	unguided planning
Accuracy	63% **	73% **	71% **	n/a	n/a
Complexity	1.23 ***	1.35 ***	1.52 ***	1.36 ***	1.62 ***
no. of pauses	37.0 ***	17.0 ***	17.0***	28.0 ***	14.0 ***
total of pauses	91 secs ***	26 secs ***	29 secs ***	41 secs ***	20 secs ***

*** = p<.01; *** = p<.001*

These discussion task scores showed that the non-native speakers in the planning conditions were able to produce language that was more fluent, more accurate and more complex than those in the no planning condition. Apart from accuracy, which was not calculated, similar effects were found for the native speakers. Without planning time the native speakers were more complex and fluent in their language than their non-native counterparts, but given planning time they too were able to increase the syntactic complexity and reduce the dysfluency of their language. Furthermore, the effects of guided planning on the non-native speakers revealed an interesting tradeoff between accuracy and complexity. With unguided planning the participants were able to increase the mean syntactic complexity of their language from 1.23 clauses per c-unit to 1.35 clauses per c-unit. With the extra help of guided planning, this increased significantly to a mean of 1.52 clauses per c-unit. However, the parallel gain in accuracy from 63% to 73% which the unguided planning had achieved did not hold up under the guided planning condition, and instead fell back 71%, a small but statistically significant difference. It appeared that the guided planners were attempting ambitious language over which they had poor control, even with the extra attention afforded by planning time.

If we now take the results of the syntactic and lexical analyses together, it is possible to discern a different tactic operating in the native and non-native speakers. Both groups had to form opinions about appropriate punishments for a list of offenders, a task deemed to be cognitively fairly demanding. Under the pressure of having to produce language on this

tricky subject without time to plan its content, the native speakers used a higher proportion of lexicalised language (much of which was not propositional but functioned as discourse organisers and fillers), a narrower range of lexicalised sequences, a moderate level of syntactic complexity and some degree of pausing. When planning time eased this pressure, there was a corresponding fall in the proportion of lexicalised language, from one in three words to one in four. The lexicalised phrases they did use were considerably more varied and there was much less need for time-filling phrases such as *'I don't know' 'I mean' 'or something' 'sort of'* as well as less need to pause while speaking. In addition to these differences, the planners were also using language that was much more syntactically complex. Taken together, these observations suggest that planning time enabled the native speaker planners to use a more fluent, open-choice, rule-based style of language than their non-planning fellow native speakers.

For the non-native speakers however, planning had no impact at all on the level of lexicalised language, which remained much lower than that detected in the native speaker data. The use of lexicalised language therefore does not seem to be a time-gaining strategy for non-native speakers. For them, the greatest time-gaining strategy seems to be the most obvious, and that is: pausing. In the unplanned condition, they paused often (a mean of 37 times) and at length (a mean total of 91 seconds out of the five minutes.) When planning time was available pausing was dramatically reduced (to a mean of 17 times and a mean total of 28 seconds) while syntactic complexity and accuracy increased. This suggests the non-native speakers were using a rule-based approach to language production which requires either pausing or, better, planning time to execute. With the help of guided planning, we see ambition in rule-based production outstripping ability to maintain accuracy.

Hypothesis 3 therefore is upheld for the native speakers only: increased attention was associated with fewer lexicalised sequences. This result fits the model of the brain as a limited-capacity information processor, and indicates that ready-made language is a useful speech production strategy to deal with heavy demands on limited attention. Hypothesis 3

was not supported for the non-native data and suggests that non-native speakers may not have a sufficient store of ready-made language which they can exploit strategically when attentional demands are great. An alternative explanation for the low levels of lexicalised sequences found in the non-native speaker data is that these speakers may in fact have a store of memorised sequences of words which are strategically useful, but not native-like and not identifiable. However, the improved accuracy of the non-native speakers' language under planned conditions suggests that they are more likely to be composing their language from rules, than drawing it from memory. The use of idiosyncratic memorised sequences of words, and an over-reliance on rules to formulate utterances are both second language production strategies that teachers might need to discourage. The former might foster ungrammatical fluency (i.e. pidgin), while the latter rests upon the often unjustified notion that what is grammatical is also acceptable. This is unfortunately not so. The classroom implications of these research findings will be discussed in full in chapter nine below.

Chapter Nine

Concluding Discussion

In this final chapter we shall consider how the results obtained from the present research study can be related to the research findings which were reviewed in the first chapters. It is interesting that, since this study was begun in 1995, several others with similar research questions (concerning non-native speech only) have been completed, and therefore we need also to discuss how the results which they report fit into the emerging picture. In addition we shall consider issues of how research of this nature is carried out, and what refinements might be made. We shall also address the relevance of this body of research to pedagogy, before finally considering useful directions for future investigations into the nature of planning and language performance.

9.1 Planning time influences native speaker performance

To date no other empirical research studies have compared the features of planned and unplanned speech in native speakers, although the speculations of Ochs (1979) are cited, usually briefly, by most researchers interested in the planned and unplanned speech of *non-*native speakers. The results reported here offer some confirmation for these speculative ideas, though the full picture is not a simple one. There is, for example, support for Ochs' ideas that complex syntax requires planning time to execute, and that word repetition and replacement are reflections of insufficient attention and thus are more frequent in unplanned than planned speech. Perhaps surprisingly, Ochs did not include pausing, reformulating or false starting in her speculative list of the features of unplanned speech, but the data analysis in this study indicates that these were numerically, if not statistically, more common in the no planning condition.

Ochs did not consider how the nature of the speaking task itself would be reflected in the language used to perform it, something clearly demonstrated in the present study. The results of the complexity analysis, for example, showed that planning time had little effect on the task that was cognitively relatively undemanding (describing the route to your home), presumably because its familiar content absorbed little in the way of attention and allowed sufficient to be directed towards language form even without planning. Although planning time offered greater attentional capacity, this was not used to complexify the language chosen to describe the route, which remained appropriately (and considerately) *un*complex. The cognitively more taxing narrative and decision tasks provided greater scope for planning, and this was exploited by the native speakers to produce more complex propositions, suitable to the subject matter. This suggests that native speakers' attention can be overtaxed by the combination of relatively complex subject matter and the desire to say relatively complex things about it, but '*undertaxing*' attention, so to speak, by relatively simple subject matter does not mean that spare capacity will be directed towards complexifying syntax. Relatively simple syntax need not necessarily be described as a processing strategy for situations of communicative stress.

Additionally, it is clear that complex syntax is not always the result of cognitive effort. The results of the lexical analysis of the decision task have indicated that native speakers use lexicalised sequences to ease the language processing load, producing more of them when attentional resources are scarce. Fixed sequences of words impose no syntactic processing burden beyond that necessary to connect them to what precedes and follows. Their internal structure can remain unanalysed and, to this extent, capacity-free. Thus, a sequence such as '*be that as it may*' appears to be syntactically complex (and indeed would have been counted as such in the complexity analysis of this data) but is actually no more complex than a single word. Sequences which have variable elements impose only the burden of providing the items necessary to complete the intended meaning. Thus '*I'm sorry to keep you waiting*' demands capacity to produce an appropriate noun phrase, tensed main verb and object pronoun, but does not demand equal capacity to produce the

infinitive *to keep* after *sorry*, or *wait* + ing after *to keep* because these are produced unanalysed[33] . Again, a complexity analysis such as was performed on the data from this study would not distinguish between subordinated clauses that are part of a lexicalised sequence, and those which are not, instead counting the two kinds as equally capacity-demanding, which is clearly not the case.

It is interesting therefore, that for the decision task the planning condition was associated with fewer lexicalised phrases (and presumably fewer instances of unanalysed subordinate clauses) but with more subordination. It follows that subordinated clauses in the planning condition were less likely to have been of the lexicalised and capacity-*un*demanding variety. If we consider that some of the credit for complexity in the no-planning condition was thus undeserved because it was due to unanalysed 'ready-to-use' language, then the increase in subordination in the planning condition reflects an even greater increase in attention to language form than would otherwise appear to be the case. Native speakers with time to plan what they were going to say about a relatively complex subject relied less on language that was simple and/or ready-to-use, and produced more that was complex and/or tailor-made.

One final observation on the native speaker performance is that their grammatical accuracy was not compromised by the pressure of having to speaking extemporaneously, even on subject matter that was considered to be cognitively taxing. That is to say, in the native speaker corpus as a whole there were no instances of incorrect verb morphology, no incorrect singular-plural concord, no mistaken articles, and no wrong use of pronoun gender, case or number.[34] It appears therefore that in terms of information processing such knowledge of grammar is proceduralised and unaffected by attentional demands. Moreover, of the total of nineteen errors detected in the native speaker corpus, seventeen were lexical in nature, e.g. *she took his knife* for *she took his life*, *warring sanctions* for

---

[33] Of course, for someone, such as a telephone help-line operator, who constantly uses the phrase, 'I'm sorry to keep you waiting' this would be a single-choice, invariable and unanalysed sequence.

warring factions, enlisted for *conscripted* and the interesting blend of two similar sequences as in *take into mind the fact that*. These were evenly distributed among the planned and unplanned conditions, and indicate (rare) failures in lexical retrieval, not in grammatical knowledge. What seems to require attentional capacity is the putting together of clauses into syntactically complete and complex combinations, the keeping track of where you are in an utterance.[35] The native speakers doing the narrative and decision tasks without planning time kept track of where they were by formulating relatively unsubordinated utterances, interspersed with pauses. Those with planning time on these tasks were able successfully to complete more complex and fluent strings.

9.2.1 Planning influences non-native speaker fluency

As mentioned above, several other studies into the effects of planning on non-native speaker performance have been completed recently. These are Ting (1996) Wendel (1997) Skehan and Foster (1997) Wigglesworth (1997a) Mehnert (1998), and Ortega (1999). All operationalise planning in the same way as in the present study, namely, the availability of a certain amount of time immediately before the performance of an experimental task, and all (with the exception of Ting) investigated oral language. The results do not figure in the literature review of the chapter three because they were not available at the time the research hypotheses of this study were being formulated. However, it is important now to consider how they relate to the findings reported here, and published as Foster and Skehan (1996)

Considering first planning time and fluency, all these studies measured fluency in different ways: through words per utterance and unpruned speech (Ortega 1999), number of words and number of T-units (Ting 1996), number and length of pauses (Skehan and Foster

---

[34] There were a few forms which could have been classified prescriptively as grammatical errors, such as *hisself* as a reflexive pronoun, or the use of *don't* for the third person, but these are more properly dialect features, and were not counted as errors.

[35] This is something Deputy Prime Minister John Prescott is famously unable to do when speaking publicly and passionately. Whether this feature of his speech is not apparent in his private, everyday conversations is an interesting question, but beyond the scope of this present investigation

1997) number of syllables per minute and mean length of pause (Wendell 1997), number of self-repairs and a type/token ration (Wigglesworth 1997a), pruned and unpruned speech rate, mean length of run, total pause time and number of pauses (Mehnert 1998). Clearly, there is no consensus as to what measure best captures the notion of fluency in speech production. But with the one exception of Ting's (rather odd) choice of the number of T-units, all these measures were found to favour the planned condition over the non-planned condition. This is a reliable indication that, whatever way you look at it, the extra attention afforded by planning time enables non-native speakers to produce more fluent language. We can conclude that the native speaker results reported here are entirely consistent with this body of research, and that dysfluency is not solely an outcome of speaking in an L2 while under some processing pressure. It is the case in L1 speech also. However, the non-native speaker results of this study suggested (rather than demonstrated) that extra planning time might lead to more repetitions and repairs as speakers with planning time were able better to monitor what they were saying. This is an interesting suggestion, and one which the most popular measures of speech rate and length /frequency of pausing have not raised. The broader brush chosen by this study, by using an extensive range of measures of dysfluency, allows a more detailed picture to emerge of how fluency and attention interact in L2 speech.

9.2.2. Planning time influences non-native speaker syntactic complexity

If we consider how syntactic complexity is calculated in these recent studies, we see a similar disparate choice of measures. There are words per utterance (Ortega 1999), number of complex constructions per narrative (Wendel 1997, Wigglesworth 1997a), clauses per c-unit (Skehan and Foster 1997) and words per c-unit, clauses per T-unit, and s-nodes per T-unit (Mehnert 1998). There seems to be little desire among SLA researchers either to agree on the most reliable measure, or (with the notable exception of Mehnert) to measure their data in several ways to enable cross-study comparisons, an unsatisfactory state of affairs to which we shall return below. However, all these studies find that, for one or more of the measures they employ, planning time is associated with significantly greater syntactic complexity. Again, the results obtained here for both native

and non-native speakers is consistent with this general picture. Speakers with time to plan what they are going to say can produce longer and more complex units, whether these be classified intonationally (utterances) or syntactically (c-units, T-units, or AS units). The native speaker dimension added by the present study is useful in showing that there are processing constraints on oral language performance even with the benefit of ideal (i.e. native speaker) grammatical competence. However, as the analysis in chapter eight had indicated, non-native speakers are less likely to be able to exploit lexicalised sequences to ease the burden of bottom-up syntactic processing which consequently has to consume more attention.

9.2.3 Planning time influences non-native speaker accuracy

Concerning accuracy, the results of recent studies are conflicting. For Ting (1996) Wendel (1997) and Ortega (1999) planning time is not associated with any increase in grammatical accuracy, only with improvements in complexity and fluency. This is consistent with Crookes (1989) and Ortega (1995a). For Foster and Skehan (1996), Wigglesworth (1997a) Skehan and Foster (1997) and Mehnert (1998) planning time does significantly increase accuracy and complexity (though not for all tasks and all measures). This is consistent with Ellis (1989). Attempting to account for her results, Wigglesworth (1997a) suggests that the testing situation, under which her study took place, provided the necessary environment for a sharper focus on language form. Her subjects were influenced by their goal in performing the task, i.e. to speak as accurately as possible in order to get a good test score. Because other studies did not take place in a testing environment, subject were not so much concerned with being accurate. Foster and Skehan (1996) and Skehan and Foster (1997) find strong confirmation for a model of information processing in which planning time allows greater attention to be paid to accuracy as well as complexity. Wendel (1997), on the other hand, accounts for the absence of a significant planning effect on accuracy in his study by distinguishing between off-line and on-line processes. He argues that pretask planning is a conscious, off-line strategic process which cannot promote accuracy because this depends on largely automatic, largely unconscious on-line monitoring. Only complexity and fluency can be affected by off-line processing. Ortega

(1999), who also found no effect for planning on accuracy, is sympathetic to this reasoning. She used post-task interviews to try to illuminate what her subjects were focussed upon during the planning time, and concludes that an individual's natural predisposition towards either communication or accuracy in L2 performance is a key factor in determining which is prioritised, and that more of her learners were consciously oriented towards prioritising communication.[36]

Of course, results are only as good as the measures used to obtain them, and at the bottom of this debate lies the question of how accuracy in language performance should be calculated. It is perhaps significant that the studies which show the most robustly postive effect on accuracy for planning time are those (Foster and Skehan 1996, Skehan and Foster 1997, Mehnert 1998) which use a global measure, whereas those which fail to show any effect (Crookes 1989, Ortega 1995a, 1998 and Wendel 1997) are those which have chosen specific measures such as the target-like use of third person -s, plural -s, the definite and indefinite articles, and/or verb morphology. Wigglesworth (1997) who did report some effect for planning on accuracy using specific measures, reports nevertheless no effect on plural -s, a rather limited effect on the target-like use of the indefinite article (two out of four tasks) and an even more limited effect on verb morphology (one out of four tasks, high proficiency learners only).

Foster and Skehan (1996) and Skehan and Foster (1997) rejected the idea that improvements in accuracy could be detected by choosing to look only at specific measures, especially ones such as plural -s, third person -s, past tense -ed and the article system which are (mostly) communicatively redundant as well as phonologically unstressed (i.e particularly difficult for the transcriber to hear). There is no good reason to suppose that an improvement in accuracy will not be evenly distributed and that these forms in particular will be used more accurately while others may not. Restricting the

---

[36] Ting (1996), whose study of Chinese L2 learners attempted to replicate Ellis (1987) without the confound of speech and writing unfortunately introduced the new confound of phonological and graphological L2 knowledge. The lack of an effect for accuracy in this study might be explained by the subjects being more proficient phonologically than graphologically.

measure of accuracy to improvements in specific forms seems wholly unsatisfactory. But global measures can be unsatisfactory also unless carefully defined. Foster and Skehan (1996) and Skehan and Foster (1997) adopted error-free clauses as an improvement on Crooke's (1989) global measure of the error-free T-unit on the grounds that a T-unit could extend over several dependent clauses. Having to reject as inaccurate a whole T-unit because of one error (which is presumably what Crookes had to do, and which may account for the lack of a significant result) gives a much smaller chance of detecting changes in accuracy than an analysis based on the (usually shorter) clause. Mehnert (1998) took this argument one stage further by noting that a clause-based analysis of accuracy makes no distinction between clauses that contain one error and clauses that contain several. She therefore supplemented a calculation of error-free clauses with a calculation of errors per 100 words. While Foster and Skehan (1996) and Skehan and Foster (1997) both find a significant effect for planning on error-free clauses in non-native speaker performance, Mehnert's results just fail to reach significance ($p = .051$) on this measure. However, her calculation of errors per 100 words does show a significant difference in favour of the planning condition ($p = .018$) while her specific measures (word order and lexical error) are far from significant ($p = .315$ and .420).

Ortega (1999) argues against global measures of accuracy as they are 'too broad to capture small changes in target-like use because they conflate multiple error types and obscure errors in grammatical domains that may be important at a given level of development' (p 118). This, however, introduces the tricky questions of what level of development the experimental subjects are at, whether they are all at precisely this level, and what grammatical features ought to be focussed upon, and whether the L1 makes a difference, and if so what this difference might be, none of which have uncontroversial answers (pace Pienemann 1984). Ortega does go on to say that possibly the best course of action is for research into accuracy to employ both global and specific measures, as Mehnert (1998) in fact does, and with interesting results, but Ortega (1999) disappointingly does not. On the balance of evidence, it seems more probable that changes in accuracy in non-native speaker performance can be captured by global measures but

missed by specific measures, but while there remains a lack of agreement among researchers on this topic, a two-pronged approach using both types should perhaps be adopted in future. Meanwhile, it would be very interesting to know what a global measure of accuracy would turn up if applied to Crookes (1989) Ortega (1995a and 1999) and Wendel (1997)[37] . The debate about whether levels of L2 accuracy can be manipulated through planning time might then be resolved.

It has been important to dwell on this question at some length because one of the central conclusions of the research reported in Foster and Skehan (1996) and Skehan and Foster (1997) is that accuracy and complexity in L2 performance are both positively affected by planning time, but that when planning is ambitious they enter into competition with each other which ultimately leads to a trade-off: complexity is prioritised at the expense of accuracy. Ortega (1999) disputes this and argues that while planning time can increase a focus on form, this is manifested (contra-intuitively) in complexity and fluency only. Wendel (1997) agrees that accuracy is unaffected by planning time, proposing post-hoc that planning is a conscious off-line strategy that cannot affect the largely unconscious process of monitoring speech on-line. In this regard, Mehnert's (1998) results are particularly illuminating. She investigated the effects of different lengths of planning time (one, five and ten minutes) on German L2 speaking tasks. She found that the one-minute planners were significantly more accurate and fluent than the nonplanners, the five-minute planners were more fluent but not more accurate than the one-minute planners, and the ten-minute planners were more fluent than the five-minute planners, though again they were not more accurate. (In fact on one of the two tasks accuracy levels actually declined.) Levels of complexity were unchanged by one or five minutes of planning, but showed a significant increase after ten-minutes. In other words, fluency increased with different lengths of planning time in an almost linear fashion, accuracy increased after only one minute of planning time but did not increase further with five or ten minutes, and complexity was only increased after ten minutes of planning time. Mehnert argues that her

---

[37] Strictly speaking, these four studies can only claim to have found no effect for planning on the accuracy of the items they chose to look at. None of the studies attempts to justify extrapolating from these items in

results show accuracy as the first priority of planning time, with complexity benefitting only after accuracy has been attended to, while fluency increases in tandem with planning time. She believes that her subjects in the ten-minute planning condition used the extra time given to them to allocate resources to complexity at the expense of further improvements to accuracy. One might add that gains in accuracy might anyway be limited by the extent of existing grammatical knowledge; learners unaware (consciously or unconsciously) of morphological and syntactic rules will not be able to access them, with or without longer planning time. In dissecting planning time in this way, Mehnert has thrown an interesting light upon it and provided further evidence that attention is a limited processing resource which is taxed by the demands of L2 production and which furthermore is selectively allocated in predictable ways. Her finding that accuracy is improved by a very short opportunity to plan: a) confirms Wigglesworth (1997a), b) supports the interpretation offered here that complexity increases with planning time while accuracy does not, and c) and indicates that comparing the results of different lengths of planning time might be a useful way to test whether and how accuracy in L2 performance is affected by attentional demands.

9.3 The interaction of planning and task type has a significant and predictable influence on L2 performance

It is notable that task type has not figured as an independent variable in most research into the impact of planning time. (Though certainly it has elsewhere, as discussed in 5.3 above). With the exception of the present study and Mehnert (1998), most other studies have used video or drawings for story-retellings or description, (Ellis 1987, Ortega 1995a 1999, Ting 1996, Wendel 1997). One other has used tasks that involve the giving of instructions or directions (Crookes 1989). These choices of task type seem to have has been governed by what would be most likely to elicit monologic speech for analysis, rather than what cognitive demands the task would itself involve. Foster and Skehan (1996) and Mehnert (1998) however have shown that this consideration should not be ignored. As

---

particular to accuracy in general.

demonstrated by the native speaker results from the present study, different tasks can impose different attentional loads even on speakers with proceduralised grammatical knowledge that takes no capacity to execute. Mehnert compared non-native speaker performance on two tasks. One was very similar to the personal task employed in the present study. Speakers had to explain on an answering machine the route a friend should take to come from Heathrow to the university in London where the research was located. This task was hypothesised to be relatively easy i.e. to rely on familiar information, to have a very obvious structure (i.e. the beginning, middle and end of the route) and to require the use only of the present and future tenses. The other task was to leave a message on an answering machine apologising to the same friend for failing to meet them at the university as arranged in the previous message, and giving reasons why. This was considered to be cognitively more taxing because it did not rely on familiar information, did not have a predictable structure and required the use of the relatively more difficult past tense. Although Mehnert's results were rather mixed, she found task type and planning time did interact. The relatively untaxing 'directions' task caused the subjects to prioritise fluency and accuracy over complexity, while the relatively more taxing 'apology' caused the subjects to concentrate on expressing complex ideas at the expense of accuracy and fluency.

Prompted by the results reported in Foster and Skehan (1996), Skehan and Foster (1997) undertook a further investigation of planning time and task performance in which the three tasks used were reasoned to be similar to the ones used here. In the Skehan and Foster (1997) study, which used the same subject pool, the personal information exchange task required subject dyads to tell each other what they found most surprising about life in England, the narrative required them to explain to their partner the story depicted in a Sempe cartoon of five or six frames, and the decision-making task required them to discuss with each other what advice to give to the writers of several 'agony aunt' letters. The results confirmed the strong effect for planning reported in Foster and Skehan (1996) on accuracy, complexity and fluency, and also confirmed the trade-off between complexity and accuracy in which the more complex the language attempted by those in the planned

condition, the weaker their control over accuracy. However, when the results fron the two studies were compared, it was clear that the task characteristics did not play out as hypothesised, that is to say that the two personal tasks did not operate in the same way, nor did the two narratives or the two decision-making tasks. The two tasks for which planning produced the greatest gains in accuracy were the 'oven' task from the present study and the Sempe cartoon, i.e. a personal information task and a narrative. The two tasks for which planning produced the greatest gains in complexity were 'find a story-line' from the present study and the 'agony aunt' i.e. a narrative and decision-making.

Skehan and Foster (1997) account for this pattern of results by hypothesising that the degree of inherent structure in a task, the degree of on-line computation it imposes and the complexity of outcomes which it requires are all important influences on the language produced by the learners who must transact it. The 'oven' task and the Sempe cartoon, though classified as different task types in the research designs both contain clearly structured subject matter: the route to a well-known place, with an obvious beginning middle and end, and an obvious story-line laid out in succeeding cartoon frames. Neither task required elements to be transformed, imagined or argued. On the other hand, the 'find a story-line' task from Foster and Skehan (1996) and the 'agony aunt' from Skehan and Foster (1997) did not contain any degree of inherent structure. The 'agony aunt' task contained complex problems with no obvious solutions, requiring a considerable degree of reflection and argument. The 'find a story-line' task did not immediately suggest any link between the pictures. These had to be imagined and then explained. Skehan and Foster (1997) conclude that the inherent and obvious task structure which can be discerned in the 'oven' task and the Sempe cartoon eases the language processing burden of the learner who can direct planning towards achieving accuracy because there are fewer competing calls on his/her attentional resources. Tasks such as the 'find a story-line' and 'agony aunt', by contrast, involve a significant degree of on-line computation and/or complex outcomes which push learners into seeking to express complex ideas, leaving less spare capacity to attend to the accuracy of their language.

This post-hoc explanation of the pattern of results reported in Foster and Skehan (1996) and Skehan and Foster (1997) has been followed by a further study into task structure and processing conditions (Skehan and Foster 1999). In this two narrative tasks were investigated, one with a clearly structured story-line and one which presented a very unpredictable series of events. It was shown by the results that the degree of structure was only a significant influence on the accuracy of the language produced during the task when interacting with some degree of planning. When planning was allowed, the structured narrative was associated with much greater accuracy than the unstructured narrative (64% error-free clauses vs 40% error-free clauses). This is interestingly consistent with Mehnert (1998) who, as we noted above, found evidence for learners using planning time to prioritise accuracy during the relatively structured 'directions' task compared to prioritising complexity during the relatively unstructured 'apology' task. It was also shown in Skehan and Foster (1999) that the structured narrative generated consistently more fluent language no matter what the planning condition, suggesting that the sequential nature of the storyline itself had aided fluency.

When taken together with other research into task type and cognitive difficulty, reviewed above in chapter 5, we can see that task characteristics are proving to be a fruitful area for research. Investigations into task type done since the present study was conceived are starting to show that the categorization used here of personal, narrative and decision-making was an interesting starting point, but far too crude. Finer distinctions are now possible, and should encourage and inform further research into the impact of particular task characteristics on L2 performance. This clearly has implications for L2 pedagogy, which is the subject matter of the section that follows.

9.4 Implications for language testing

The analysis of the both the native and non-native data in the present study has suggested that for all language users there is a discernible tension between language form and language meaning. For native speakers this tension seems to be generated by the burden of

organising complete and complex syntactic strings to express propositions on taxing or unfamiliar subjects. For non-native speakers this tension is greater because of the added burden of having simultaneously to process capacity-demanding L2 morphosyntax and lexis. This is an important and hitherto ignored consideration for most L2 oral tests. Well-known tests of EFL oral proficiency, such as the UCLES[38] and Trinity exams, expect candidates to show their proficiency in English by speaking with native-like phonology, fluently and accurately, and with appropriate complexity on an unprepared topic. No allowance is made for the possibility that even someone with native proficiency might be too taxed by the task to be able to display such a tour de force.[39] The test designers do not consider, for example, that the complexity and unfamiliarity of a topic such as describing the differences between six wristwatches, speculating on the lifestyle of the inhabitants of a particular house, or discussing the care of the elderly[40] , might in themselves cause problems for candidates, pre-disposing them to produce rather more dysfluent, inaccurate or inappropriately simple language than would be the case if they had prepared for the topic and were therefore less burdened by its content.

Wigglesworth (1997a) has shown that even very short periods of planning time (one or two minutes) can result in learners producing language that is significantly more fluent, complex and accurate. Mehnert (1998) has shown that accuracy is improved by only one minute of planning time. Consequently, giving candidates a brief period to prepare what they are going to say could permit them to display the higher, rather than the lower, limits of their proficiency. Anecdotally, I can remember preparing my own EFL students for oral exams by making them learn by heart the sort of lexicalised sequences that would allow them to appear to be fluent and accurate if thrown into confusion by a tricky task (*Well, I wonder what on earth could be going on in this photograph. Yes, that is a very interesting question, but you have to look at it from both sides,).* This tactic may have been

---

[38] University of Cambridge Local Examinations Syndicate.

[39] It can be added that some L1 speakers seem to be naturally more hesitant than others, not through any linguistic deficit, but because they appear to be more apt to change their minds in mid-utterance, or to have some trouble choosing the right words for what they wish to say. (Jan Hulstijn, personal communication)

[40] All of these tasks appeared in the Cambridge Advanced English oral exams in 1993

successful (though as an oral examiner I could spot it immediately) and with perplexing tasks it may have been necessary in order to disguise mute desperation. But it meant candidates were not being given a proper chance to show what they could do creatively with English, which ought to have been the object of the exercise.

The native speaker results reported here should persuade L2 oral examiners not to assume that dysfluency, syntactic inaccuracy or simplicity are necessarily signs of poor proficiency, but may be caused by the task itself. Native speaker piloting of tasks could help to weed out those which cause problems even for speakers with effortless language skills. Additionally, offering candidates just a few minutes pre-task planning time might enable testers to reach a better and fairer assessment.

9.5 Implications for teaching

If we can see planning time as influencing the quality of a learner's linguistic output, then it is possible to speculate that this influence might extend beyond performance in tests (*putting on your best show*) to promoting progress in second language acquisition *(advancing in target language ability)*. Because planned output is more complex and (at least according to the results reported here) more accurate, it can be seen as pushing interlanguage to its limits, thereby engaging acquisitional processes and making it more likely that interlanguage forms move in the direction of the target language. Similarly, it can be argued that the more often language forms at this 'cutting edge' of interlanguage are accessed and used, the more likely it is that such forms will shift from planned, thoughtful use to more spontaneous control. In the terms used in cognitive psychology research reviewed in chapter one, planning gives opportunities for practice, which is important in transforming a skill from being procedural and capacity-demanding to being automatic and capacity-free. In this view pre-task planning time actively facilitates interlanguage development, and is therefore a useful classroom activity.

9.5.1 Planning and a focus on form

As discussed at length in chapter five above, there is currently a debate in the field of second language pedagogy on the best way to engineer focus on language *form* i.e. how to attract a learner's attention to formal aspects of the target language within the context of meaningful language use. To be avoided is any intervention, such as a prolonged exposition of language rules (*forms*), which necessarily diverts the central focus of attention away from meaning. However, many of the suggested techniques involve the teacher selecting the forms for attention, through for example the phonological or graphical highlighting of problem morphemes in a text (Long and Robinson 1998). This, as we noted, is unsatisfactory because it means the teacher is choosing where to direct everyone's attention, and this may not be appropriate for all learners in the class. According to Ortega (1999), from a focus-on-form perspective, planning time is theoretically promising, not only because it lightens the cognitive load and allows attention to be directed away from meaning and towards form during the task performance itself, but also because during the pre-task planning stage it can cause the learner to attend consciously to formal aspects of the language which he or she realises are necessary to accomplish the task ahead. What these formal aspects might be, and how closely to attend to them, depends entirely on the individual learners as they assess the linguistic demands of the task. This kind of focus-on-form, initiated and regulated by the learners themselves, could be important in causing learners to become aware, at timely moments, of the 'gaps' (Schmidt 1994) or 'holes' (Swain 1998) in their L2 knowledge. Ortega's (1999) research using retrospective interviews with Spanish L2 learners on what they were thinking about during pre-task planning shows that while some were devoting most of their attention to formal aspects of the language, others were in fact prioritising meaning, indicating that individual differences in approach to pre-task planning are an important factor in determining how far form will be focussed upon when it is the learners themselves who choose where to allocate their attention. Nevertheless, it can be claimed that giving planning time before an L2 task creates opportunities for learners to attend to the language forms they need to use during the task itself, and that these 'learner-generated' insights are likely to be more valuable than any focus on specific forms contrived by the

teacher. The research results reported here would support such a view, though clearly only a longitudinal study would be able to show whether frequent opportunities to engage in pre-task planning is a significant independent variable in interlanguage development.

Related to this discussion is the question, explored in chapter three, of whether languages are most effectively learned through implicit-inductive learning, or explicit-deductive learning, that is to say whether grammatical patterns in the target language take root in the learners' mind thanks to their own conscious or subconscious analyses, or are planted there by teachers and text-books. A review of the research evidence concluded that for complex systems such as natural languages, (as opposed to simple letter strings and light sequences) implicit-inductive learning through mere exposure is not nearly as effective as explicit-deductive teaching and/or focussed exposure to helpful examples. (e.g.Ellis 1993, DeKeyser 1995) However, advocates of a focus-on-form approach (e.g. Long and Robinson 1998) take the view that explicit-deductive teaching needs to be kept to a minimum because explicit focussing on language forms can actually impede second language acquisition by encouraging the learner to rely too much on declarative knowledge of L2 rules.

The research study reported here was not designed to illuminate this debate, which will surely continue for some time yet. However, in the discussion of the results so far, it has been argued that planning time eases the cognitive load of an L2 task and consequently allows learners to access, analyse and ultimately advance their knowledge of the L2 either during the planning or performance of the task. The question arises therefore of where this knowledge comes from in the first place. Should the task participants be noticing patterns or form-function mappings in the task input material itself (Prabhu 1987)? Alternatively, should they be having their attention drawn momentarily to certain forms selected by the teacher from the task input or the learner output (Long and Robinson 1998)? Or should they be using the task in order to map functions to particular language forms pre-selected by the task designer (Samuda in press)? Or should they perhaps be merely gaining experience in manipulating particular language forms which have just been presented and explained at length by a PPP teacher (Harmer 1991)? The methodology chosen to engage

learners with the L2 forms will inevitably depend on how implicit or explicit the teacher believes language learning to be. But what language forms are actually 'noticed' or focussed upon during task planning and task performance are, I would argue, too unpredictable to be worth the effort of control. A focus on form or on forms can take place in spite of the teaching methodology as well as because of it. What *is* sure is that no noticing, focussing on or learning of L2 forms can happen in the absence of examples, and ultimately perhaps it does not make much difference whether these examples are the ones which have been: a) helpfully included by the task designer, or b) assumed to be essential to the successful transaction of a particular task, or c) momentarily highlighted by the teacher with a marker pen or exaggerated intonation, or d) explicitly taught as part of a structural syllabus, or e) remembered consciously by the learners as being similar to something else encountered before in the input, or f) noticed subconsciously as part of a new pattern. Pre-task planning time increases the chances that learners will be able to draw upon L2 knowledge, *from whatever source*, and use it to increase the complexity, accuracy and fluency of their interlanguage performance. In other words, pre-task planning time is a useful technique that can be grafted onto a variety of methodologies. It would be equally at home in a classroom organised by Long and Robinson (1998) as one organised by Harmer (1991).

9.5.2 Native speaker examples as a focus on form

It was argued in section 5.1 above that of all the suggested techniques for focussing on form, the most promising was recasting in that it provided timely native speaker examples for comparison with problematic interlanguage utterances, and allowed the learner to notice any points of comparison between the two. This is useful in demonstrating to an alert learner the morphological, lexical and syntactic gaps in his or her interlanguage knowledge, and can be used at any moment when the learner is interacting with a teacher. However, as we have seen from the discussion on chapter eight, words do not necessarily pattern according to syntactic rules alone, but to idiomatic conventions. It is common for learners to produce L2 utterances which are grammatical but decidedly odd. It would be very interesting to know whether teachers are less likely to recast these than they are to

recast utterances which contain some grammatical error. It is possible that teachers have a higher tolerance for the unidiomatic than for the ungrammatical.

It was shown in the lexical analysis of the data reported in chapter eight that the non-native speakers used a much lower proportion of lexicalised sequences than did the native speakers, and that, in contrast to the native speaker data, this proportion was unaffected by planning time. A tentative conclusion from studying this data was that the non-native speakers were constructing a great proportion of their language from rules rather than lexicalised routines, with or without the benefit of planning time. One might therefore speculate that classroom teaching which valued accuracy more than fluency would encourage this rule-based strategy with the unfortunate result that the learners are helped to be grammatical at the cost of being slow and odd. However, as we noted in chapter eight, it is possible that the non-native speakers in this study were in fact using idiosyncratic memorised sequences, which no-one else could detect. If it is the case that language learners build their own memory bank of fixed or partially fixed sequences that may not be native-like, then there is a concomitant danger that successfully and regularly used examples will become too deeply entrenched in memory to be dislodged easily by further reflection. In the terminology of the experimental psychology research reviewed in chapter two, these routine sequences might be on course to becoming automaticised, difficult to suppress once established in long-term memory, and very hard to replace with another re-analysed sequence (Shiffrin and Schneider 1977). Inappropriate but useful sequences of words are, in this sense, undetectable signs that the learners are creating their own idioms, and these could be one source of fossilised error.

To help prevent learners from committing inappropriate word sequences to memory, and/or to encourage them to build a more native-like memory store, it may be useful for teachers to enable learners to reflect upon their own language use and to compare it to native-speaker norms in a more systematic way than is offered by occasional recasting. This most definitely does not mean devoting part of the class time to an 'idiom of the day'. In the past there has perhaps been too much of an assumption that teaching idiomatic language means phrase-book type parroting, and no-one wants to learn by heart, and out

of context, phrases which they might never use appropriately. Far better would be giving learners the chance to see what language choices native speakers make in specific contexts. A task-based classroom could be an ideal place for this. Learners who have done or who are about to do a language task could listen to a tape or read a transcript of native speakers doing exactly the same task. (cf. COBUILD course, Willis and Willis 1988) Individual learners could thereby be prompted into reanalysing some part of their own interlanguage sequences or adding some new native speaker ones which they notice as useful. Additionally, teachers could choose to draw the explicit attention of the whole class to some native speakers routines. Again, the choice of method depends upon how implicit-inductive or explicit-deductive the teacher wishes to be.

To take an example from the discussion task transcripts used in this study, the non-native speakers were fond of a variety of expressions centered around the verb 'to agree'. It appeared many times and in many forms, some grammatical and some not.

I agree with you
I don't agree with you
I am agree with you
I am not agree with you
I am agree.
Are you agree with me.?

In the native speaker data, the verb *agree* occurs rarely and in a rather different way. (*So, we're pretty much agreed on this one, I agree with that*). For the native speakers, expressing agreement was usually done by using the phrases *That's right* or *that's true* expressing disagreement was usually done rather obliquely by beginning an utterance with *then again*, or *yeah but.* The non-native speaker who overuses the grammatical *I agree with you* and *I don't agree with you* needs to expand his/her repetoire to become more native-like. The learner who constantly uses the ungrammatical *I am not agree with you* and *are you agree with me* needs to to reanalyse this verb as well as perhaps replace it in a discussion type context with a more native-like selection, an outcome which the teacher

might explicitly or implicitly facilitate. However, it is fair to say that the teacher is not always the best source for information on appropriate native-like selections. Much of what is native-like in language remains below the level of conscious awarenesss[41] and is revealed only by extensive corpus analysis (Sinclair 1991). Presenting learners with native speaker examples of the target language, as used in exactly the same tasks as they are being given to do, allows them to notice particular patterns of use which perhaps the teacher would not recognise as worth noticing

Thus the use of native speaker examples can have an important function in prompting leaners to become aware of differences between the ways they would express themselves in a particular context and the ways native speakers do. It can also help to provide them with appropriate lexicalised sequences of words that ease the processing burden of composing speech. Building a memory store of lexicalised sequences that are subjected to analysis and reanalysis is a way for learners to become more fluent, more accurate *and* more target-like.

9.6.1 Future directions for research into planning and L2 performance

This research study was undertaken to explore part of Skehan's (1996) theoretical framework for task-based language learning. The framework described how task design and task implementation could serve to manipulate learners' attention between the competing goals of accuracy, complexity and fluency. Many of the lines of inquiry suggested in this have already been followed. For example, Skehan and Foster (1997) and Foster and Skehan (in preparation) have looked at how a learner's attention can be manipulated through post-task activities as opposed to pre-task planning. Foster and Skehan (1997) investigated how pre-task planning could be disrupted by mid-task intervention. On the subject of the influence of task type on performance, Skehan and Foster (1999) contrasted the outcomes of two kinds of narrative, one more structured

---

[41] The fact that the native speakers in the data preferred to agree or disagree with 'that' rather than with 'you' was something which surprised me. I had never noticed this before, but having been alerted, I now

in its storyline than the other. The results gained so far are indicating that mid-task intervention has little effect, the post-task activity of learner's transcribing their own recorded performance has some positive effect upon levels of accuracy, and task type appears to interact with planning time to give selective improvements in particular performance areas.

Taken together with the results reported here, one of the most fruitful lines for future research would seem to be this interaction of planning time with task type. There is scope for more research into the selective effects of task characteristics on the nature of performance, and results obtained in such an inquiry would be of practical use in the classroom. So far the indications are that tasks with a clear structure, well-known subject matter and no planning time are likely to lead learners to prioritise fluency in their performance (Foster and Skehan 1996, Skehan and Foster 1997): that tasks leading to complex outcomes, or those where information has to be transformed or interpreted, are likely to lead learners to prioritise complexity (Skehan and Foster 1997): and in tasks where planning is allowed on familiar, structured information learners are likely to prioritise accuracy (Foster and Skehan 1996). Although some other researchers have reported results which are consistent with these interpretations, (Mehnert 1998, Bygate 1996a) it is nevertheless very important that other studies be undertaken that attempt to replicate these findings. It is unfortunate that much second language acquisition research seeks to innovate rather than replicate because this often leads to widespread acceptance of research conclusions that are based on a very few small-scale studies. (The negotiation of meaning, discussed in chapter three, is a case in point). Future research might also usefully consider whether the effects for planning reported here in the non-native speaker data hold up in longitudinal and within-subjects designs. It is possible that the cross-sectional snapshot taken by the present research (and, it has to be said, all other published research into the effects of planning) is not an accurate prediction of what effects long term pre-task planning has on interlanguage

---

notice it in native-speaker usage all the time.

development[42] Within-subjects designs were used by Ellis (1987), Crookes (1989), Ortega (1995a and 1999) and Ting (1996) and have shown positive effects for planning. Therefore, the untried combination of longitudinal and within-subjects design would be a very useful basis for investigation, with potentially interesting applications for pedagogy.

Another line of research is to investigate different operationalisations of planning. Following Crookes (1989) the study reported here chose to use ten minutes of silent reflection before the task began, with the added refinement of providing half of the non-native planners with a sheet of paper detailing some useful ideas on how to use the time.[43] This proved a useful distinction in that detailed planning promoted the greatest complexity, while undetailed planning lead to an emphasis on accuracy. Clearly, further investigation is warranted. Foster and Skehan (in press) have already looked at whether planning time is most effectively orchestrated by the teacher or by learners in groups, and whether it is possible to direct learners' attention selectively towards language content or language form. Preliminary results suggest that planning time orchestrated by the teacher results in more complex *and* accurate performance, and that learners preparing in groups did not perform nearly as well, though in neither set-up was it possible to control the focus of learners' attention between form and content. This raises the interesting question of whether teacher input serves best to sharpen rather than direct learners' attention. Further studies could also set out to investigate the interactions between length of planning time, type of planning condition and characteristics of task design with the ultimate aim, perhaps, of providing teachers and syllabus designers with a sound empirical underpinning for the selection and

---

[42] Tracking the non-native speakers across three weeks of data gathering, which was done in the present study, gives it a somewhat longitudinal dimension, but clearly this is very limited.

[43] Ortega (1999) reveals that Crookes (1989) actually guided his planners with suggestions on how to plan during the ten minutes, making the detailed planning condition of the study reported here very similar to the single planning condition used by Crookes. This is a nice illustration of how important it is for researchers to give full details of their methods so that replications or comparisons can be undertaken. However, it does not, as claimed by Ortega, call into question any claims made for the effects of planning type on performance by this and related studies.

implementation of classroom tasks best suited to particular pedagogic goals, whether this be to increase fluency, promote interlanguage development through greater complexity, or consolidate L2 knowledge through greater accuracy.

9.6.2 The need to establish more reliable measures in research

As I have tried to make clear in the discussion of the results for accuracy in section 9.2.3 a major problem for research of this type is the variety of measures used to assess such aspects of L2 performance as syntactic complexity, fluency and grammatical accuracy. The problem is compounded by the general reluctance many second language acquisition researchers exhibit when it comes to explaining in detail exactly what definitions they applied to such measures as the 'utterance' or 'the 's-node'. Foster, Tonkyn and Wigglesworth (forthcoming) note that it is common for research reports not to include any definition at all of the measures used, as if assuming that everyone in the readership will know what these measures are and therefore no explanation is necessary. This is unfortunately not true. Not only is there no consensus on how best to assess aspects of L2 performance, there are no widely available definitions of the measures most commonly used. Thus the *'s-node'* of one research study may be identical in definition to the *'clause'* of another, or it may be different to the *'s-node'* used in a third. Some researchers feel that the issue of definition of terms is not as important as a high interrater reliability score for individual studies (Ortega 1999), but I would argue that without an accepted and accessible definition of terms it is not possible to generalise from research findings or carry out robust cross-study comparisons, both of which procedures are necessary in any field of scientific inquiry.

This is a highly unsatisfactory state of affairs that future research needs to address. To a great extent Malvern and Richards (1997) have tackled the shortcomings of the type-token ration, (the most popular measure of lexical density), and have proposed a promising new measure based on a much more sound mathematical model. Foster, Tonkyn and Wigglesworth (forthcoming) offer a detailed and comprehensive definition of a syntactically-based unit to measure spoken interaction which, it is hoped, will

provide a reliable and accessible tool for research analyses on both L1 and L2 transcripts. However, the issues of measuring fluency and accuracy remain to be tackled, and in the absence of a consensus, researchers should perhaps be willing to produce figures for a range of measures for the same variable. If nothing else, this will provide interesting information on how different measures perform. Its main benefit, however, would be to allow a more reliable picture to emerge from a comparison of results from different sources, an outcome devoutly to be wished.

9.7 The role of native speaker studies in future L2 research

Unremarkably perhaps, research into second language acquisition has either totally ignored native speakers, or very occasionally has allowed them the role only of interlocutors for the non-native speakers under investigation (for a typical example see Pica et al 1989). There was perhaps an assumption that native speakers were useful as models of perfect target language competence, or else as testers of non-native speaker communicative competence, but nothing more. The native speaker data here has shown that what we ask learners to do in tasks is not just dependent on language proficiency, but also on the extent to which the task content is known or understood. More specifically, it has shown that dysfluent, syntactically simple utterances can be the result of uncertainty in the face of difficult subject matter, and not just inadequate target language proficiency. Research into task types therefore ought ideally to include native speaker base-line data so that assumptions made about why non-native speakers performed the way they did in experimental tasks can first be tested against the performance of speakers with native proficiency. This approach would better enable researchers to understand what aspects of task design tax the parts of the brain not connected to language processing. Native-speaker piloting would be especially useful in designing tasks for language tests in which a candidates' inability to deal with complex or perplexing material, as opposed to complex and perplexing language, should not be part of the assessment.

The lexical analysis of the data , reported in chapter eight, was a preliminary attempt to use native-speaker intuition to identify the degree to which lexicalised sequences were used by both native and non-native speakers. The results were interesting, and suggest that this analysis could usefully be applied to the other two tasks. Equally, it would be helpful to know how individuals differ in their use of lexicalised sequences under different task and planning conditions, something not possible with the present research design which could only compare group scores. Although the identification of lexicalised sequences entailed a laborious coding process by several native speakers, and is thus not lightly to be undertaken, it provided valuable data that could not be collected in any other way. Any research which seeks to understand the strategic role of memory in native and non-native speech production could employ this approach to obtain quantitative measurements. It would be fascinating, for example, to know the relationship between L2 proficiency (as measured by professional oral language examiners) and use of native-like lexicalised sequences (as counted by native speaker intuition).

9.8. Finally.

In a footnote to her 1999 paper, Ortega says that the issue for research into how language learners allocate attention 'is not whether learners *do* attend to form and meaning concurrently, but whether they *can*'. (Her emphasis). I would argue that the opposite is far more important and practical. We need to know what learners *are likely to do*, so that we can design classroom procedures that seek to harness instincts useful to SLA. To do this properly we need to bear in mind what native speakers are likely to do with language tasks, remembering that learners are, for the most part, ordinary people with ordinary human instincts.

Appendix A

Personal task: instructions to students (undetailed planning condition)

Sending someone back to turn off the oven!

Instructions

It is the afternoon and you have an important exam in fifteen minutes. You suddenly think that you haven't turned off the oven after cooking your lunch. There is no time for you to go home.

Explain to a friend who wants to help:
- how to get to your house
- how to get into your house and get to the kitchen
- how to turn off the oven

You have ten minutes to prepare for the task.
You can make notes during the ten minutes but you won't be allowed to use these notes while doing the task.

Make you notes below. Use the other side of the paper if you need more space.

Name:_____

Appendix B
Narrative task: instructions to students (detailed planning condition)

Picture story

Instructions

In this task you have to tell a story from a series of pictures.

You have ten minutes to prepare for this task

Don't think about only one thing in each picture: try to make your story include as much as possible from each picture.

You can make notes during the ten minutes, but you won't be allowed to use these notes while doing the task.

These are things you can do to help you to prepare:

- think about how the pictures can be made into a story.
- think what connects the different pictures.

- plan the things you want to emphasise during the story.
- think about what problems your listener could have, and how you might help her.
- think about how your listener can understand the order of the story.

- think about what grammar you need to do the task.
- think about what vocabulary you need to do the task.
- think how to avoid difficulties and solve problems with grammar and vocabulary.

Make you notes below. Use the other side of the paper if you need more space.

Name:_____

Appendix C
Decision making task: instructions to students (no planning condition)

Judge task

Instructions

You and your partners are judges. You have four decisions to make. For each decision you must decide how long to send the accused to prison. The maximum is a real life sentence. The minimum is three months. Or you may set them free.

1. The accused found her husband in bed with another woman. She took the breadknife and killed him.

2. The accused is a prisoner of war. Your country has just defeated his. He was a pilot and dropped an atom bomb on your tenth largest city, killing 200,000 people and injuring many more.

3. The accused is a doctor. He gave an overdose (a very high quantity) of a pain-killing drug to an 85-year-old woman because she was dying painfully of cancer. The woman had asked for the overdose. The woman's family accuse the doctor of murder.

4a) Three teenage boys were having a fight with a fourth boy near a swimming pool. They threw him into the water and then stood on him till he drowned.

4b) Five adults were sitting near the pool and watched the fight. They did nothing to help the fourth boy.

Appendix D: Typical problem cases in oral data segmentation[44]

Problem A) 'Because' adverbial clauses

Optional adverbial clauses introduced by 'because', are frequently problematic in oral language. This is due to the fact that the relationship between the 'because' clause and its putative main clause is often uncertain. 'Because' often performs a discourse marker function. In examples 1 and 2 below "because" could be paraphrased as 'I say this because...'. This discourse function is frequently signalled by pause and intonation phenomena, or is used as a way of continuing after an interlocutor has responded to an utterance

1. A. I think she should leave him↘ (0.5) <u>because I had a boyfriend like that once</u>.

2. A. I think I would ask er some- no er a judge↘.
   B. Yes!
   A. <u>because she is quite in dangerous position.</u>

To be counted as subordinate in an AS unit analysis, the adverbial clause must be final <u>and</u> within the same tone unit as at least one of the other preceding clause elements of the AS-unit. This linkage establishes that the adverbial clause is definitely part of the plan which produced the initial main clause. Otherwise it is counted as a syntactically separate unit. In example 1 the falling intonation (↘) and pause mean that '*because I had a boyfriend like that once'* is a new AS unit. The falling intonation and pause for B's interjected *'yes!'* in example 2 means that '*because she is in a dangerous position'* is also counted as a new AS unit.

Problem B) Co-ordination

T-unit analysis decrees that co-ordinated verb phrases should be treated as belonging to one unit. In the following data it is clear from pause and intonation phenomena, that the second verb phrase constitutes a new start:

---

[44] All examples are taken from Foster, Tonkyn and Wigglesworth (forthcoming). Some are originally from the data collected for this study

3. and the woman, um, (0.5) speaks um, um, go telephone, speaks telephone, (1.0) and look in in the other direction,

4. The other woman is very happy now (0.5) and (3.0) just walking away with a gr- great smile.

In many instances, non-native speakers will drop subjects in this way, creating apparent co-ordinated phrases which are almost certainly not intended as such. Therefore in cases where co-ordination of verb phrases occurs, an AS unit analysis will consider the co-ordinated phrases as belonging to the same AS-unit, except in cases where the first phrase is marked by falling or rising intonation and is followed by a pause of at least 0.5 seconds. In examples 3 and 4 above therefore, *'and look in in the other direction'* is counted a separate AS unit, as is *'and (3.0) just walking away with a great smile'.*

Problem C) 'Topical' noun phrases

As Bygate (1988) has noted, independent noun phrase satellite units are common in speech, and particularly so in the case of second language learners whose first language is typologically a 'topic-comment' language Neither the T-unit, nor the c-unit provides a satisfactory answer to the analysis of such phenomena. If words per unit are being measured, it is important to clarify where topical noun phrases belong. The problem is exemplified in the underlined sections of 5 and 6 below.

5 and some children they are playing the ball

6. three people, maybe they are children, three people they are swimming now.

7. It's just a matter of passing . And especially the basic education,↘ (0.5) they have to pass automatically from one grade to another.

AS-unit analysis treats topicalised noun phrases as generally belonging to the unit of which they are the topic; thus the following comment includes the (generally pronominalised) noun phrase repetition of the topic. In example 5 above, *'Some children'*

is counted as the topic and therefore as part of the following unit. In example 6 '*Three people.... three people they are swimming now'* is one AS unit with a repeated topic. '*Maybe they are children'* is another AS unit.

However, noun phrase satellite units which are separated from the following AS unit by falling intonation and a pause (equal to or greater than 0.5 sec) will be treated as separate AS-units. Thus the underlined section of example 7 above would be treated as a separate AS-unit because the word *'education'* is marked by falling intonation (↘) and followed by a pause of 0.5 seconds.

Problem D) Scaffolding and interruption

Co-operative building up of a conversation, with sharing of units, is common in highly interactive conversations. However, the analysis of such interactions raises quite complex questions about how the resulting units are to be divided up and assigned. Example 7 and 8 are typical.

8. B I think um (3.0) yeah it's hard maybe he like her but no like a girlfriend, just

A Just like a sister.

B. Yeah.

9. A: oh that's a big problem=

B: oh no!

A: = because my shop's policy is only to give the credits for the return goods

In example 8, "*just like a sister*" completes the AS-unit of speaker B, and therefore Speaker B is credited with the unit. This is also credited as a complete AS-unit to speaker A. In example 9 , A's utterances would be analysed as a single AS-unit with 3 subordinate clauses despite the interruption. This is because the previously stated criteria for including a final adverbial in the preceding AS-unit have been met. (See Problem A above).

Bibliography

Anderson, J. (1983) *The architecture of cognition.* Cambridge, MA: Harvard University Press

Anderson, J. (1982). 'Acquisition of cognitive skill'. *Psychological Review Vol 89, no 4* 369- 406.

Anderson, J. (1990). *Cognitive psychology and its implications.* Third edition. New York: Freeman.

Anderson, J. (1992) 'Automaticity and the ACT* theory'. *American Journal of Psychology.* 105 (2), 165-180.

Anderson, J. (1995) *Learning and Memory: an integrated approach* New York, Wiley

Anderson, J. and Finchum, J. (1994) 'Acquisition of procedural skills from examples.' *Journal of Experimental Psychology: Learning, Memory and Cognition.* 20 (6) 1322-1340.

Aston, G. (1986). 'Troubleshooting interaction with learners: the more the merrier?' *Applied Linguistics* 7, 128-143.

Becker, J.D. (1975) The phrasal lexicon. Report no.3081. Advanced Research Projects Agency of the Department of Defense.

Bley-Vroman, R. and Chaudron, C. (1994). Elicited imitation as a measure of second-language competence. In E. Tarone, S.M. Gass and A. D. Cohen (eds.) *Research Methodology in Second-Language Acquisition* Hillsdale, NJ: L. Erlsbaum.

Bock, K. (1982) 'Toward a cognitive psychology of syntax.' *Psychological Review* 89.1 1-47

Braine, M. (1971) 'The acquisition of language in infant and child'. In C. Reed (ed.) *The Learning of Language.* Appleton-Century-Crofts.

Brooks, N. (1960) *Language and Language Learning.* New York: Harcourt, Brace and World.

Brown, G. (1995). *Speakers, Listeners and Communication.* Cambridge: Cambridge University Press.

Brown, G, Malmkjaer and Williams, J (1984) *Performance and Competence in Second Language Acquisition* Cambridge University Press

Brown, R. (1991) Groupwork, task difference and second language acquisition *Applied Linguistics* 11 1-12

Brumfit, A. and Windeatt, S. (1985) *Communicative Grammar* English Language Teaching for the Arab World: Oxford University Press.

Butterworth, G, and Hatch, E. (1978). 'A Spanish-speaking adolescent's acquisition of English syntax,' in Hatch, E. (ed). *Second Language Acquisition* Rowley, Mass: Newbury House.

Bygate, M. (1988). 'Units of oral expression and language learning in small group interaction'. *Applied Linguistics* 9: 59-82.

Bygate, M. (1996a) 'Effects of task repetition: appraising the developing language of learners' in J. Willis and D. Willis (eds.).

Candlin, C. (1987). 'Towards task-based language learning' in C. Candlin and D. Murphy (eds.): *Language Learning Tasks.* Englewood Cliffs, N.J.: Prentice Hall.

Cheng, P.W. (1985). 'Restructuring versus automaticity: alternative accounts of skill acquisition'. *Psychological Review, Vol 92, no 3.* 414- 423

Chomsky. N. (1957). *Syntactic Structures.* The Hague: Mouton.

Chomsky. N. (1965). *Aspects of the Theory of Syntax.* M.I.T. Press.

Corder, S. (1967). 'The significance of learners' errors'. *International Review of Applied Linguistics* 5. 1-20.

Corder, S. (1971). 'Idiosyncratic dialects and error analysis'. *International Review of Applied Linguistics* 9 149-59.

Corder, S. (1974). 'Error analysis' in J. Allen and S. Corder (eds). *The Edinburgh Course in Applied Linguistics,* Vol 3. Oxford: Oxford University Press

Crookes, G. (1989). 'Planning and interlanguage variation'. *Studies in Second Language Acquisition* 11:367-83.

Curran, T. and Keele, S. (1993). 'Attention and structure in sequence learning.' *Journal of Experimental Psychology: Learning, Memory and Cognition.* 19, 189-202.

De Groot, A. (1965). *Thought and choice in chess.* The Hague: Mouton.

Deese, J. (1980) 'Pauses, prosody and the demands of production in language'. In W. Dechert and M. Raupach. (eds.) *Temporal variables in speech*. The Hague. Mouton.

DeKeyser, R. (1995) 'Learning second language grammar rules: an experiment with a miniature linguistic system.' *Studies in Second Language Acquisition*. 17, 3, 379-410.

Doughty, C. and Pica, T. (1986). 'Information gap tasks: Do they facilitate second language acquisition?' *TESOL Quarterly* 20: 305-25.

Doughty, C and Varela, E (1998) Communicative focus on form: p114-138 in xxxxxxxxxx

Doughty, C and Williams, J (1998) (eds*) Focus on form in classroom second language acquisitio*n. New York: Cambridge University Press

Dulay, H and Burt, M. (1974) 'You can't learn without goofing' in Richards (ed.) *Error Analysis.* London: Longman.

Dulay, H. and Burt, M. (1973) 'Should we teach children syntax?' *Language Learning* 23: 245-58

Ellis, N. (1993). 'Rules and instances in foreign language learning: interactions of explicit and implicit knowledge.' *European Journal of Cognitive Psychology*, 5, 289-318.

Ellis, N. (1994). 'Psychological perspectives on the role of conscious processes in vocabulary acquisition'. *Aila Review 11. 37- 56.*

Ellis, N. (1996). 'Sequencing in SLA: phonological memory, chunking, and points of order'.. 18 91-126.

Ellis, N. and Large, B. (1987). 'The development of reading: As you seek so shall you find'. *British Journal of Psychology*. 78. 1-28.

Ellis, R. (1986) *Understanding Second Language Acquisition* Oxford: Oxford University Press.

Ellis, R. (1987). 'Interlanguage variability in narrative discourse: style shifting in the use of the past tense'. *Studies in Second Language Acquisition* 9:12-20.

Ellis, R. (1993) 'The structural syllabus and second language acquisition.' *TESOL Quarterly*, 27, 91-113.

Fantoni, B. (1985) *Colemanballs.* Private Eye/Andre Deutsche

Favreau, M. (1981). *Automatic and Conscious Attentional Processes in First and Second Languages of Fluent Bilinguals: Implications for Reading.* unpublished doctoral dissertation

Felix, S. (1980) 'Interference, interlanguage and related issues'. In Felix, S. (ed.) *Second Language Development.* Tubingen: Gunther Narr.

Firth, J, R. (1935) 'The technique of semantics.' *Transactions of the Philological Society,* 36 -72.

Fisk, A,. and Schneider, W. (1983). 'Category and word search: Generalising search principles to complex processing.' *Journal of Experimental Psychology: Learning, Memory and Cognition, 9,* 177- 195.

Foster, P. (1998) 'A classroom perspective on the negotiation of meaning'. *Applied Linguistics.* 19.1.

Foster, P. and Skehan, P. (1996). The influence of planning on performance in task-based learning. *Studies in Second Language Acquisition,* 18, 3, 299-324.

Foster, P. and Skehan, P (1997) Modifying the task: the effects of surprise, time and planning type on task-based foreign-language instruction. *Thames Valley University Working Papers in English Language Teaching. Vol.4.*

Foster, P. and Skehan, P. (ms) 'The influence of source of planning and focus of planning on task-based performance'. Centre for Applied Linguistic Research. Thames Valley University.

Foster, P. and Skehan, P. (in press). The effects of post-task activities on task-based performance.

Foster, P. Tonkyn, A and Wigglesworth, G (forthcoming) Measuring spoken language: A unit for all reasons.

Gass, S and Varonis, E. (1985). 'Task variation and non-native/non native negotiation of meaning'.in S. M. Gass and C.G. Madden (eds). *Input and Second Language Acquisition.* 149-62. Rowley, MA: Newbury House.

Gathercole, S. Willis, C. Baddeley, A. and Emslie, H. (1994) 'The children's test of non-word repetition: a test of phonological working memory'. *Memory.* 2 103-127

George, H. (1972) *Common Eerrors in Language Learning: Insights from English.* Rowley Massuchusetts: Newbury House.

Givon, T., (1979) 'From discourse to syntax: grammar as a processing strategy.' In T. Givon (ed.) *Syntax and Semantics, Vol 12, Discourse and Syntax.*(pp 81-113) New York: Academic Press.

Goldman-Eisler, F. (1968) *Psycholinguistics: experiments in spontaneous speech.* New York: Academic Press.

Grauberg, W. (1971) An error analysis in the German of first year university students. In Perren, G. and Trim J.(eds.) *Applications of Linguistics.* Cambridge: Cambridge University Press.

Gregg, K. (1984) 'Krashen's Monitor and Occam's razor' *.Applied Linguistics* 5. 79-100.

Harmer, J. (1991). *The Practice of English Language Teaching.* London: Longman.

Hatch, E. Polin, P. and Part, S. (1970) 'Acoustic scanning or syntactic processing'. Paper presented at a meeting of the Western Psychological Association. San Francisco.

Hirst,W., Spelke,E., Reaves, C., Caharack, G., and Neisser, U. (1980). 'Dividing attention without alternation or automaticity'. *Journal of Experimental Psychology: General, 109*, 98- 117.

Hoffman, J., Nelson, B., and Houck, M. (1983). 'The role of attentional resources in automatic detection'. *Cognitive Psychology, 15*, 379- 410.

Howarth, P. (1998) Phraseology and second language proficiency. *Applied Linguistics* 19.1 24-44

Hulstijn, J. (1992) Implicit and incidental second language learning: Experiments in the processing of natural and partly artificial input. In H.W. Dechert and M. Raupach (eds.). *Interlingual Processes.* Tubingen: Gunter Narr Verlag.

Jackson, H. (1981) 'Contrastive analysis as a predictor of errors, with reference to Punjabi learners of English'. In J. Fisiak (ed.) *Contrastive Linguistics and the Language Teacher.* Oxford, Pergamon.

James, C. (1980) *Contrastive Analysis* London: Longman

Jefferson, G. (1997) 'Preliminary notes on a possible metric which provides for a 'standard maximum' silence of approximately one second in conversation.' In Roger, D. and Bull, P. (eds) *Conversation: An Interdisciplinary Perspective.* Multilingual Matters. Clevedon. Philadelphia.

Karmiloff-Smith, A.(1986) From meta-processes to conscious access: evidence from children's metalinguistic and repair data. *Cognition* 23 95-147

Kolers. P., and Roediger, H., (1984). 'Procedures of mind'. *Journal of Verbal Learning and Verbal Behavior 23,* 425- 449.

Krashen, S.D. (1981). *Second Language Acquisition and Second Language Learning.* Oxford. Pergamon.

Krashen, S.D. (1982). *Principles and Practice in Second Language Acquisition.* Oxford: Pergamon.

Krashen, S.D. (1985). *The Input Hypothesis.* London/New York: Longman

Krashen. S.D. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *Modern Language Journal,* 73, 440-464.

Krashen, S.D., Butler, J., Birnbaum, R., and Robertson, J. (1987). 'Two studies in language acquisition and language learning'. *ITL: Review of Applied Linguistics* 73-92.

Kroll, B., (1977). 'Ways communicators encode propositions in written and spoken English: A look at subordination and co-ordination.' in E. Keenan and T. Bennett, (eds.) *Discourse Across Time and Space,* SCOPIL, University of Southern California, Los Angeles, California.

LaBerge, D. and Samuels, S. (1974). 'Towards a theory of automatic information processing in reading'. *Cognitive Psychology, 6,* 293-323.

Lado, R. (1957) *Linguistics across cultures: Applied linguistics for language teachers.* Ann Arbor, Michigan. University of Michigan

Lee, W. (1968) 'Thoughts on contrastive linguistics in the context of language teaching.' In J Alatis (ed.) *Contrastive Linguistics and its Pedagogical Implications.* Washington. D.C. Georgetown University.

Lenneberg, E. (1967). *Biological Foundations of Language.* Wiley.

Levelt, W., (1978). 'Skill theory and language teaching.' *Studies in Second Language Acquisition 1*, 52-70.

Levelt, W. (1989) *Speaking: from intention to articulation* Cambridge University Press

Lewis, M. (1993). *The Lexical Approach*. Language Teaching Publications

Long, M. (1985). 'Input and second language acquisition theory' in S.Gass and C. Madden (eds.): *Input in Second Language Acquisition*. Rowley, Mass: Newbury House. pp377-93.

Long, M. (1989). 'Task, group, and task-group interaction'. *University of Hawaii Working Papers in English as a Second Language* 8: 1-26.

Long, M. and Crookes, G. (1991). 'Three approaches to task-based syllabus design'. TESOL Quarterly 26: 251 86.

Long M and Robinson, P. (1998) Focus on form: theory, research and practice. In Doughty, C and Williams, J (1998) (eds*) Focus on form in classroom second language acquisitio*n. New York: Cambridge University Press

Lyster, R. (1998) Recast, repetition and ambiguity. *Studies in Second Language Acquisition* 20.1 51-81

Lyster, R and Ranta, L. (1997). Corrective feedback and learner uptake: negotiation of form in communicative classrooms. *Studies in Second Language Acquisition* 19,37-66.

Malvern, D. and Richards, B. (1997). 'A new measure of lexical density'. In Ryan, A. and Wray, A. (eds.) *Evolving Models of Language*. Clevedon: Multilingual Matters.

McLaughlin, B. (1987) *Theories of Second Language Acquisition*. London: Edward Arnold.

McLaughlin. B., Rossman, T. and McLeod, B. (19xx) 'Second language learning: an information-processing perspective'. *Language Learning*.33.2 135-158.

Mehnert, U. (1998). The effects of different lengths of planning time on second language performance. *Studies in Second Language Acquisition*. 20, 83-108

Moon, G. (in prep) Frequencies and forms of phrasal lexemes in English.

Nattinger, J. and Decarrico, J. (1992) *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.

Norusis, M. (1990) *SPSS PC User's manual, Version 4*. Chicago: Statistical Package for the Social Sciences.

Nissen, M. J. and Bullemer, P. (1987). 'Attentional requirements of learning: evidence from performance measures.' *Cognitive Psychology*, 19, 1-32.

Nunan, D. (1989). *Designing Tasks for the Communicative Classroom*. Cambridge: Cambriidge University Press.

Ochs, E., (1979) 'Planned and unplanned discourse'. In T. Givon (ed.) *Syntax and Semantics, Vol 12, Discourse and Syntax.*(pp 51-80) New York: Academic Press.

Ortega, L. (1995). The effect of planning in L2 Spanish narrative. (Research note no.15) Honolulu:University of Hawai'i, Second Language Teaching and Curriculum Centre.

Ortega, L (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*. 21.1.109-148.

Pawley, A. and Syder, F. (1983) 'Two puzzles for linguistic theory: nativelike selection and nativelike fluency'. In Richards, J. and Schmidt, R (eds) *Language and Communication.* London: Longman.

Pica, T. (1994). 'Research on negotiation: What does it reveal about second language learning conditions, processes, outcomes?' *Language Learning* 44/3: 493-527.

Pica,T Young, R and Doughty,R (1987)The impact of interaction on comprehension. *TESOL Quarterly* 21 737-58.

Pica, T., Halliday, L., Lewis, N. and Morgenthaler, L. (1989). 'Comprehensible output as an outcome of linguistic demands on the learner'. *Studies in Second Language Acquisition* 11: 63-90.

Pienemann, M (1984) Psychological constraints on the teachability of languages. *Studies in Second Language Acquisition* 6: 186-214

Pienemann, M (1989) Is language teachable? *Applied Linguistics* 10 52-79

Pinker, S. (1994). *The Language Instinct.* Allen Lane: Penguin Press

Posner, M. and Syder, C. (1975). 'Attention and cognitive control'. In R.L. Solso (ed.) *Information processing and cognition: The Loyola Symposium.* Hillsdale, N.J. Erlbaum.

Prabhu, N, (1987) *Second Language Pedagogy*. Oxford: Oxford University Press.

Raupach, M, (1984) 'Formulae in second language speech production.' in H. Dechert, D. Mohle and M.Raupach (eds.) 1984: *Second Language Productions*. Tubingen: Gunter Narr.

Reber A. (1989) 'Implicit learning and tacit knowledge'. *Journal of Experimental Psychology: General*. 118, 219-235.

Reber, A. (1976) 'Implicit learning of synthetic languages: the role of instructional set.' *Journal of Experimental Psychology: Human Learning and Memory*. 2. 88-94.

Reber, A., Kasin, S., Lewis, S. and Cantor, G. (1980). 'On the relationship between implicit and explicit modes in the learning of a complex rule structure.' *Journal of Experimental Psychology: Human Learning and Memory*.6, 492-502

Robinson, P. (1996) (ed.) *Task complexity and Second Language Syllabus Design: Data-based studied and speculations*. University of Queensland Working Papers in Language and Linguistics

Robinson, P. (in press) 'Task complexity, cognition and second language syllabus design: a triadic framework for examining task influences on SLA.'

Robinson, P., Ting, S. and Urwin, J (1996) 'Investigating second language task complexity.' *RELC Journal*, 26, 62-79.

Rivers, W.M. (1964). *The psychologist and the foreign language teacher*. Chicago. University of Chicage Press.

Rulon, K. and McCreary, J. (1986). 'Negotiation of content: teacher-fronted and small-group interaction.' in R. R. Day (ed.) *Talking to Learn*: 182-9. Rowley. MA: Newbury House.

Rumelhart, D., and Norman, D., (1978) 'Accretion, retuning and restructuring: three modes of learning'. In Cotton, J. and Klatsky, R. (eds.), *Semantic factors in cognition*. Hillsdale, N.J: Lawrence Erlbaum.

Sachs, J. (1967) 'Recognition memory for syntactic and semantic aspects of connected discourse'. *Perception and Psychophysics* 2: 437-442.

Sacks, H., Schegloff, E., and Jefferson, G. (1974) 'A Simplest Systematics for the Organisation of Turn-Taking in Conversation'. *Language* 50, 696-735.

Samuda, V.(in press) 'Guiding relationships between form and meaning during task performance: the role of the teacher' in Bygate, M .,Skehan, P, and Swain, M (eds) *Language tasks: Teaching, learning and testing*, London: Longman.

Schmidt, R (1994) 'Deconstructing consciousness in search of useful definitions for applied linguistics'. *AILA Review* 1994 11-26

Schmidt, R. (1983). 'Interaction, acculturation, and the acquisition of communicative competence'.in N. Wolfson and E. Judd, (eds). *Sociolinguistics and Second Language Acquisition*. Rowley Mass.: Newbury House.

Schmidt, R. (1990). 'The role of consciousness in second language learning'. *Applied Linguistics* 11 129-157.

Schmidt, R. (1993a). Awareness and second language acquisition. *Annual Review of Applied Linguistics*.13, 206-226.

Schmidt. R. (1994) Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review* 11. 11-26.

Schmidt, R. and Frota, S. (1986). Developing basic conversational ability in a second language: a case-study of an adult learner. In R.Day (ed.) *Talking to Learn* Rowley, Mass: Newbury House.

Schneider, W and Shiffrin, R. (1977). 'Controlled and automatic human information processing: 1. Detection, search and attention'. *Psychological Review, 84,* 1 - 66.

Schneider, W. and Detweiler, M. (1988) 'The role of practice in dual-task performance: Towards workload modelling in a connectionist/control architecture'. *Human Factors* 30 539-566.

Schumann, J. (1979). 'The acquisition of English negation by speakers of Spanish a review of the literature.' in Andersen, R. (ed.) *The Acquisition and Use of Spanish and English as First and Second Languages* Washington DC TESOL.

Schumann, J. (1980).'The acquisition of English relative clauses by second language learners'.in Scarcella, R. and Krashen, S. (eds.) *Research in Second Language Acquisition*

Schwarz, B. (1993) On explicit and negative data effecting and affecting competence and linguistic behaviour. *Studies in Second Language Acquisition.* 15, 147-163.

Scollon, R. (1976) *Conversations with a one-year-old.* Honolulu, Universtiy of Hawaii Press

Searle, J.R. (1976). A classification of illocutionary acts. *Language in Society*. 5.1-23.

Seliger, H. (1979). 'On the nature and function of language rules in language teaching.' *TESOL Quarterly* 13. 359-69.

Selinker, L. (1972). 'Interlanguage'. *International Review of Applied Linguistics.* 10. 209-30.

Sharwood-Smith, M. (1981). 'Consciousness-raising and the second language learner.' *Applied Linguistics* 7 239-256

Sharwood-Smith, M. (1993) 'Input enhancement in instructed SLA: theoretical bases.' *Studies in Second Language Acquisition* 15 165-80

Shiffrin, R. and Schneider, W. (1977). 'Controlled and automatic human information processing: 2. Perceptual learning, automatic attending, and a general theory'. 1977 *Psychological Review, 84, 2,* 127- 190.

Shortall, T. (1996) 'What learners know and what they need to know'. in Willis, J. and D. Willis (eds.) *Challenge and Change in Language Teaching*. Heineman.

Sinclair, J. (1991) *Corpus, Concordance, Collocation.* Oxford: Oxford University Press.

Skehan, P. (1986) 'Cluster analysis and the identification of learner types'. In Cook,V. (ed.) *Experimental Approaches to Second Language Acquisition*. Pergamon Press.

Skehan, P. (1989) *Individual Differences in Second Language Learning*, Edward Arnold.

Skehan, P. (1996). 'A framework for the implementation of task-based instruction'. *Applied Linguistics*, 17, 38-62

Skehan, P. (1998) *A Cognitive Approach to Language Learning.* Oxford University Press

Skehan, P and Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research* 1. 185- 211.

Skehan, P and Foster, P. (1999) The influence of task structure and processing conditions on narrative retellings. *Language Learning* 49.1.93-120

Skinner, B. (1957) *Verbal Behaviour* Appleton-Century-Crofts.

Soars, J and Soars, L (1986a) *Headway Intermediate*. Oxford University Press.

Soars, J and Soars, L (1986b) *Headway Intermediate Teacher's Book*. Oxford University Press.

Stannard Allen, W. (1947) *Living English Structure*. Longman.

Stockwell, R., and Bowen, J. (1965). *The Sounds of English and Spanish*. Chicago: Chicago University Press.

Swain, M. (1985). 'Communicative competence: Some roles of comprehensible input and output in its development.' in S. M. Gass and C.G. Madden (eds). *Input and Second Language Acquisition.* 115-32. Rowley, MA: Newbury House.

Swain, M. (1998*) 'Focus on form through conscious reflection'* in C. Doughty and J. Williams (eds.) Focus on from in classroom second language acquisition (pp 64-81). New York: Cambridge University Press.

Tarone, E. (1985). Variability in interlanguage use: A study of style-shifting in morphology and syntax. *Language Learning*, 35, 373-403.

Thomas, M. (1998). Programmatic ahistoricity in second language acquisition theory. *Studies in Second Language Acquisition* 20,3, 387-407.

Thompson. A.J. and Martinet, A.V. (1980) *A Practical English Grammar*. Oxford University Press.

Ting, S. (1996). Planning time, modality and second language task performance. *University of Queensland Working Papers in Language and Linguistics.* 1.31-64.

Todd, L. (1984) *Modern Englishes: Pidgins and Creoles.* Oxford: Basil Blackwell

Trask, R. (1995). *Language: The Basics*. Routledge.

VanPatten, B. (1990). 'Attending to form and content in the input'. *Applied Linguistics* 12 287-301

VanPatten, B. (1994) 'Evaluating the role of consciousness in second language acquisition: terms, linguistic features and research methodology'. *Aila Review 11*, 27-36

Varonis, E. and Gass, S. (1985). 'Non-native conversations: A model for negotiation of meaning.' *Applied Linguistics*. 6: 71-90.

Weinert, R. (1995) The role of formulaic language in second language acquisition: A review. *Applied Linguistics* 16.2.181-205.

Wendel, J. (1997) *Planning and second language narrative production*. Unpublished doctoral dissertation. Temple University, Philadelphia, PA.

Widdowson, H. (1989). Knowledge of language and ability for use. *Applied Linguistics* 10: 128-37

Wigglesworth, G. (1997a). An investigation of planning time and proficiency level on oral test discourse. *Language Testing*. 14.85-106.

Willis, D (1990) *The Lexical Syllabus: A New Approach to Language Teaching*. London: Collins.

Willis, D. and Willis, J (eds.) (1996) *Challenge and Change in Language Teaching*. Heineman.

Willis, J. (in prep) Between grammar and lexis: some exploratory issues in the identification and measurement of the use of lexical phrases.

Willis, J. (1996). *A Framework for Task-based Learning*. London: Longman.

Willis, J. and Willis. D. (1988) *Collins COBUILD English Course: Book 1*. London: Collins.

Wode, H. (1976) 'Developmental sequences in naturalistic L2 acquisition'. *Working Papers on Bilingualism* 11, 1-13.

Wode, H. (1978). 'The L1 vs L2 acquisition of English interrogation'. *Working Papers on Bilingualism* 15, 37-57.